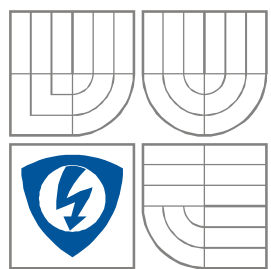


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ  
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH  
TECHNOLOGIÍ  
ÚSTAV AUTOMATIZACE A MĚŘICÍ TECHNIKY

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION  
DEPARTMENT OF CONTROL AND INSTRUMENTATION

# VÝPOČET MAPY DISPARITY Z BAREVNÝCH STEREO SNÍMKŮ

CALCULATING DISPARITY MAP FROM COLOR STEREO IMAGES

BAKALÁŘSKÁ PRÁCE  
BACHELOR'S THESIS

AUTOR PRÁCE  
AUTHOR

BARBORA KULÍKOVÁ

VEDOUCÍ PRÁCE  
SUPERVISOR

Ing. Jan Klečka

BRNO, 2016

# Bakalářská práce

bakalářský studijní obor **Automatizační a měřicí technika**

Ústav automatizace a měřicí techniky

**Studentka:** Barbora Kulíková

**ID:** 168329

**Ročník:** 3

**Akademický rok:** 2015/16

**NÁZEV TÉMATU:**

## Výpočet mapy disparity z barevných stereo snímků

### POKYNY PRO VYPRACOVÁNÍ:

Cílem práce je návrh a implementace algoritmů pro výpočet mapy disparity z rektifikovaných barevných stereo snímků.

1. Zpracujte rešerši týkající se tvorby mapy disparity jako prostředku pro vyhledávání korespondencí ve stereo páru snímků.
2. Na základě literární rešerše vyberte alespoň dva různé algoritmy, které jsou vhodné pro tvorbu mapy disparity z barevných snímků nebo je možné je k tomuto modifikovat.
3. Implementujte zvolené algoritmy v prostředí Matlab.
4. Vytvořte galerii rektifikovaných barevných stereo snímků a ověřte funkčnost Vámi implementovaných algoritmů.
5. Prezentujte dosažené výsledky, zhodnoťte klady a zápory Vašich metod.

### DOPORUČENÁ LITERATURA:

[1] SONKA M., HLAVAC V., BOYLE R.: Image Processing, Analysis and Machine Vision. 3rd edition.

Toronto : Thomson, 2008. 829 s. ISBN 978-0-495-08252-1.

[2] JAN J.: Medical image processing, reconstruction and restoration: concepts and methods. Boca

Raton: Taylor, 2006, 730 s. ISBN 08-247-5849-8.

**Termín zadání:** 8.2.2016

**Termín odevzdání:** 23.5.2016

**Vedoucí práce:** Ing. Jan Klečka

**Konzultant bakalářské práce:**

**doc. Ing. Václav Jirsík, CSc., předseda oborové rady**

### UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

## **ABSTRAKT**

Tato bakalářská práce se věnuje tvorbě disparitních map. První kapitoly jsou zaměřeny na fyziologii lidského vnímání prostoru a metodách zobrazení 3D obsahu, což jsou témata blízce související s disparitní mapou, jejím vznikem a praktickým využitím. Následuje popis používaných metod zpracování obrazu. Stěžejní teoretická kapitola pojednává o metodách výpočtu disparitní mapy a použitých principech. V praktické části byla vytvořena aplikace s uživatelským rozhraním v prostředí MATLAB. Aplikace umožňuje tvorbu disparitních map pomocí lokálních a globálních metod. Funkčnost aplikace a implementace metod je experimentálně ověřena a byl proveden experiment srovnávající jednotlivé metody a zkoumající vliv parametrů lokální metody na kvalitu disparitní mapy. Součástí práce bylo vytvoření jednoduché databáze stereosnímků.

## **KLÍČOVÁ SLOVA**

Disparity, disparitní mapa, hloubková mapa, stereoskopie, MI, SSIM, CGRAD, Depth Discontinuities by Pixel-to-Pixel Stereo, SSD

## **ABSTRACT**

This bachelor's thesis deals with a creation of the depth maps. The first chapter concerns with the physiology of a human space perception and the methods of displaying the 3D content which are the topics closely related to the depth map, its creation and practical usage. Subsequently, there is a chapter focused on the description of the used methods of the image processing. The fundamental theoretical chapter deals with the methods of computing the disparity and used principles. In the practical part of the thesis an application has been made with a user interface in the Matlab environment. The application enables the user to create the disparity maps through the local and global methods. The functionalism of the application and the implementation of the methods are experimentally verified. An experiment comparing the methods and analyzing influence of the local method parameters on the quality of the depth map has been made. The last part of the thesis was to create a simple stereopicture database.

## **KEYWORDS**

Disparity, disparity map, depth map, stereoscopy, representation of the image, MI, SSIM, CGRAD, Depth Discontinuities by Pixel-to-Pixel Stereo, SSD

Kulíková, B. *Výpočet mapy disparity z barevných stereo snímků*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Ústav automatizace a měřicí techniky, 2015. 47 s. Bakalářská práce. Vedoucí práce: ing. Jan Klečka

## PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma Výpočet mapy disparity z barevných stereo snímků jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

V Brně dne .....

.....

(podpis autora)

## PODĚKOVÁNÍ

Děkuji vedoucímu bakalářské práce ing. Janu Klečkovi za účinnou metodickou, pedagogickou a odbornou pomoc a další cenné rady při zpracování mé bakalářské práce.

V Brně dne .....

.....

(podpis autora)

# OBSAH

<b>Seznam obrázků</b>	<b>vi</b>
<b>Seznam tabulek</b>	<b>vii</b>
<b>1 Úvod</b>	<b>1</b>
<b>2 Mechanismus stereopsie a stereoskopie</b>	<b>2</b>
2.1 Mechanismus stereopsie .....	2
2.2 Akomodace oka .....	2
2.3 Srovnání pozorování 2D snímku, 3D scény a 3D světa .....	3
2.4 Metody stereoskopické projekce .....	4
2.4.1 Anaglyf .....	4
2.4.2 Pasivní 3D projekce .....	4
2.4.3 Aktivní 3D systém s časovým multiplexem .....	5
2.4.4 Autostereoskopické displeje .....	5
<b>3 Vymezení základních pojmů</b>	<b>7</b>
<b>4 Použité barevné modely a předzpracování</b>	<b>9</b>
4.1 Reprezentace obrazu .....	9
4.2 Použité metody zpracování obrazu .....	10
<b>5 Metody výpočtu DISPARITNÍ mapy</b>	<b>11</b>
5.1 Výpočet mír podobnosti.....	12
• Absolutní rozdíl (AD, z angl. absolute difference).....	12
• Rozdíl čtverců .....	12
• Vzájemná informace (MI, z angl. Mutual Information) .....	12
• SSIM – Podobnost dvou obrazů (Structural Similarity Index).....	13
• $C_{\text{GRAD}}$ .....	13
5.2 Agregace vypočtených mír .....	13
5.3 Určení výsledné nejoptimálnější disparity.....	14
5.3.1 Lokální metody .....	14
5.3.2 Globální metody .....	14

5.4	Konkrétní algoritmy.....	19
5.4.1	Depth Discontinuities by Pixel-to-Pixel Stereo .....	19
5.4.2	SSD .....	19
5.5	Využití korespondencí výrazných bodů obrazu.....	20
<b>6</b>	<b>Implementace metod</b>	<b>21</b>
6.1	Lokální metoda .....	21
6.2	Semiglobální metoda .....	25
6.3	Belief propagation.....	25
6.4	Metoda založená na nalezení výrazných bodů .....	27
<b>7</b>	<b>Výsledky</b>	<b>30</b>
7.1	Parametry lokální metody .....	30
7.2	Porovnání implementovaných metod .....	35
7.3	Ukázka výsledků.....	37
<b>8</b>	<b>Výstupy bakalářské práce</b>	<b>39</b>
8.1	Databáze stereosnímků .....	39
8.2	Uživatelské rozhraní .....	39
<b>9</b>	<b>Závěr</b>	<b>42</b>
	<b>Literatura</b>	<b>44</b>
	<b>Seznam symbolů, veličin a zkratk</b>	<b>47</b>

# SEZNAM OBRÁZKŮ

Obr. 2.1:	Princip systému s oddělením pomocí polarizace [8].	5
Obr. 2.2:	Princip autostereoskopického zobrazování [8].	6
Obr. 3.1:	Varianty vzájemného umístění kamer v prostoru.	8
Obr. 5.1:	Základní princip šíření zprávy v belief propagation [23].	17
Obr. 5.2:	Schéma algoritmu graphcut [22].	18
Obr. 6.1:	Vývojový diagram implementovaného postupu.	23
Obr. 6.2:	Vývojový diagram funkce SlideImages zajišťující výpočet míry podobnosti a jejich agregaci.	24
Obr. 6.3:	Vývojový diagram a) semiglobální metody Scanline b) Globální metody Belief Propagation.	27
Obr. 6.4	Vývojový diagram experimentální metody využívající výrazné body obrazu.	29
Obr. 7.1:	Výstupní hloubkové mapy získané během experimentu V první řádce snímek Cloth a ve druhém Mid. V prvním sloupci použití metriky SSIM a v druhém SAD.	32
Obr. 7.2:	Obrázky z databáze [32] použité v provedeném experimentu a) Levý snímek b) Pravý snímek c) pravdivá hloubková mapa. První řada shora scéna Baby, druhá řada scéna Wood, třetí řada scéna Mid a čtvrtá řada scéna Cloth. ....	33
Obr. 7.3:	Závislost kvality disparitních map získaných různými mírami podobnosti na velikosti okna použitého pro agregaci.	34
Obr. 7.4:	Závislost kvality disparitních map získaných různými mírami podobnosti na váze míry podobnosti spočítané v gradientním prostoru.	34
Obr. 7.5	Závislost kvality vypočítané disparitní mapy metodě pro různé snímky.	36
Obr. 7.6:	Vstupní levý a pravý snímek scény Tsukuba.	37
Obr. 7.7:	Pravdivá disparitní mapa scény Tsukuba.	37
Obr. 7.8	Disparitní mapy snímku Tsukuba získané lokálními metodami při nastavení různých parametrů.	38
Obr. 8.1	Panely pro nastavení a) Lokálních metod b) Globálních metod.	40
Obr. 8.2	Uživatelské okno aplikace.	41



# SEZNAM TABULEK

Tab. 7.1: Výsledky experimentu testujícího vliv použité míry podobnosti a vliv velikosti okna pro agregaci míry podobnosti.....	31
Tab. 7.2: Výsledky experimentu testujícího vliv použité míry podobnosti a vliv váhy míry podobnosti spočítané v gradientním prostoru. ....	32
Tab. 7.3: Porovnání implementovaných metod, přejaté metody a funkce prostředí MATLAB.....	36

# 1 ÚVOD

Práce se zabývá výpočtem disparitní mapy. Disparita je rozdíl horizontálních obrazových souřadnic daného objektu mezi levým a pravým obrazem téže scény. Pomocí vnitřních parametrů kamery lze z disparity dopočítat hloubku bodu scény, tedy vzdálenost bodu od snímací kamery a jeho umístění v trojrozměrném prostoru reálného světa. Stereo vidění, a tedy i výpočet disparitní mapy, je aktuální problém. Využití nachází zejména v 3D TV, ale i tvorbě modelů objektů, které se následně dají využít v mnoha průmyslových odvětvích, jako jsou stavebnictví či strojírenství.

Vstupem pro zpracování jsou dva, případně více rektifikovaných snímků, tzv. stereo snímků, což jsou snímky téže scény, pořízené dvěma fotoaparáty lišícími se polohou v horizontální ose. Druhou možností je pořízení snímků fotoaparáty v libovolné pozici, ty však následně musí být rektifikovány. Výstupem je disparitní mapa, což je snímek o stejné velikosti jako jednotlivé vstupní snímky a reprezentuje vzdálenost korespondenčních bodů v rektifikovaných souřadnicích od kamery. Existují dvě základní skupiny metod pro zpracování stereo snímků a tvorbu disparitních map – lokální a globální. Základem obou metod je hledání korespondujících bodů ve dvou snímcích. Tedy bodů, které reprezentují tentýž bod scény.

V první části semestrálního projektu je popsán samotný princip lidského vnímání prostoru, na jehož základě fungují dnešní 3D televizní zobrazovače a projektory. Z popsaného systému vychází celý princip tvorby hloubkové mapy. Další kapitola projektu se věnuje reprezentaci reality pomocí 2D obrazů a metodami jejich zpracování. Kapitola je důležitá z toho důvodu, že samotná tvorba hloubkové mapy patří mezi metody zpracování obrazu a může používat mnoho jiných metod, například segmentaci, hranové detektory, úpravy jasové stupnice či hledání významných bodů obrazu. Následující kapitola obsahuje definici důležitých základních pojmů, které souvisí s tvorbou disparitní (hloubkové) mapy. Stejně tak teoretická kapitola 5 se věnuje popisu postupu tvorby disparitních map. V kapitole je uveden základní postup, který je rozdělitelný do několika fází. Po definici jednotlivých fází následuje vysvětlení metod a postupů použitelných v jednotlivých fázích tohoto postupu. Na závěr této kapitoly je stručně popsáno několik konkrétních metod tvorby hloubkové mapy. Přestože cílem semestrálního projektu bylo nastudování dané tematiky, tak práce obsahuje i náznak praktického řešení projektu, kterému se věnuje kapitola 6. V prostředí MATLAB byl implementován základní algoritmus lokální metody tvorby hloubkových map a byl proveden experiment testující vliv použité metriky pro vyhodnocení míry podobnosti a vliv velikosti okna ve fázi agregace míry podobnosti.

V této práci se budu zabývat zástupci obou skupin metod. Výsledkem bude program pro výpočet disparitní, případně i hloubkové mapy s příjemným uživatelským prostředím a s možností volby použité metody a nastavením parametrů. Následně budu zkoumat účinnost jednotlivých metod. Program bude vytvořen v programovém prostředí Matlab.

## 2 MECHANISMUS STEREOPSIE A STEREOSKOPIE

Každé oko člověka bez postižení zraku snímá jeden obraz. Zrak pracuje simultánně, kdy oba získané obrazy se spojí v jeden a člověk je schopen vnímat hloubku prostoru. Tomuto jevu se říká binokulární vidění [1]. Jelikož jsou oči umístěny v určité vzdálenosti od sebe, obrazy jsou posunuty a jsou i mírně odlišné.

### 2.1 Mechanismus stereopsie

Při správném zaostření se bod A zobrazí na sítnici levého oka jako bod  $A_L$  a na sítnici pravého oka jako bod  $A_R$ . Při správné vergenci obrazy bodu A na obou sítnicích splynou. Vergence označuje takové pohyby očí, které vedou k zaměření zrakových os obou očí na sledovaný objekt. Následná senzorká fúze umožňuje sloučit oba sítnicové obrazy do jednoho sítnicového vjemu [2]. Díky znalosti stavu vergence, a tím i paprsků procházejících body  $A_L$  a  $A_R$ , je mozek schopen si domyslet místo protnutí. Na tomto místě vnímáme pozici bodu A.

Bod B, který je v jiné vzdálenosti než bod A, se na sítnici nezobrazí jako bod, ale jako rozostřená skvrna. Oči jsou natočeny k bodu A, jehož obraz s obrazem B nesplývá, ale body mají jistou disparitu, tj. míru splývání. Pokud je tato disparita malá, je mozek schopen obraz domyslet a ze znalosti této disparity určit vzdálenost bodu B. Pokud je velká, neschopnost mozku ji zpracovat se projeví jako dvojité vidění. Velikost disparity záleží především na natočení očí [3].

### 2.2 Akomodace oka

Akomodace je proces, při němž se mění optická mohutnost oka, které tak může automaticky zaostřovat na různě vzdálené předměty [4]. Pokud se oči dívají na vzdálený bod, je úhel optických os očí malý a oči akomodují na nekonečno. Při dívání se na blízký bod je úhel os velký a oči akomodují na blízko. Akomodace a vergence jsou spolu úzce spjaty. Při absenci procesu vergence, tj. při sledování obrazu pouze jedním okem, by proces zaostření z blízkého bodu na vzdálený a naopak trval déle.

Proces akomodace probíhá na oční čočce a z části vysvětluje lidské prostorové vnímání. Dopadá-li na čočku rovnoběžný svazek paprsků, čočka mění jejich směr a za čočkou se paprsky protnou. Bod, který je na ose čočky, se nazývá ohnisko. Vzdálenost průsečíku od čočky je ohnisková vzdálenost. Tento pojem se používá, pokud je objektiv nahrazen jedinou tenkou čočkou. Předmětová vzdálenost je vzdálenost předmětu od čočky. Obrazovou vzdáleností se nazývá vzdálenost čočky nahrazující objektiv od roviny snímáče.

Na sítnici by měly být zaostřeny pouze body v jediné předmětové vzdálenosti od čočky. V praxi se ale ukázalo, že člověk vidí ostře předměty v jistém rozsahu předmětových vzdáleností. Tento rozsah se nazývá hloubka ostrosti [3].

## 2.3 Srovnání pozorování 2D snímku, 3D scény a 3D světa

V rámci fyziologie stereoskopického vidění existují dva typy prostorových vodítek, a to monokulární a binokulární. Monokulární se projeví při pozorování scény i jen jedním okem. Mezi monokulární vodítka patří:

- Interpozice,
- lineární perspektiva,
- atmosférická perspektiva,
- pohybová paralaxa,
- relativní výška,
- relativní velikost,
- gradient – spád struktury,
- akomodace[5].

Binokulární vodítka jsou dvě, a to:

- Binokulární konvergence, kdy je zpracovávána informace o tenzi vnějších očních svalů.
- Binokulární disparita, což je úhlový rozdíl projekcí jednoho bodu na sítnice obou očí [6].

Sledování 2D fotografie je pro diváka pohodlné, jelikož se osy očí protínají v rovině snímku, kam oči ostří, a tedy akomodace a vergence jsou v souladu.

Při pozorování reálné 3D scény se oči kříží ve vergenční vzdálenosti, kam také ostří. Obrazy v této vzdálenosti mají nulovou disparitu. Mozek z těchto informací následně odhadne vzdálenost pozorovaného předmětu. Obrazy bodu v jiné vzdálenosti budou mít nenulovou disparitu. Pokud bod ležel v oblasti ostrosti, mozek disparitní obrazy sloučí a opět vzdálenost odhadne.

Jestliže bod leží mimo oblast ostrosti, disparita je příliš velká a obrazy jsou rozmazané. Mozek nedokáže vzdálenost odhadnout a dochází k tzv. diplopii, což je jev dvojitého vidění. Pokud mozek toto dvojité vidění zaregistruje, změnou vergence sníží disparitu a změnou akomodace zvýší ostrost pro jinou vzdálenost od očí, čímž diplopii redukuje. Mozek tak získá potřebné informace o hloubce jiné části scény, kterou teď vidí perfektně. Takto člověk vnímá svět, přičemž výsledný vjem je ostrý a prostorový obraz.

Pokud je ve 3D snímku, který pozorujeme, bod s nulovou snímkovou disparitou, vergenční a akomodační vzdálenost je rovna vzdálenosti očí od plátna a tento bod tedy vnímáme v rovině plátna. Pokud se ve 3D snímku nachází bod s nenulovou snímkovou disparitou, která je dostatečně malá a nemusí tak nastat změna akomodace a vergence, mozek tuto disparitu interpretuje jako vzdálenost rozdílnou od vzdálenosti plátna. I při větší disparitě je možná interpretace bez změny obrazového vjemu. Změní se vergence očí, sníží se sítnicová disparita takřka na nulu a také se změní akomodace očí. Oči tak budou ostřit jinam, než na rovinu plátna, ale na obrazovém vjemu se to neprojeví.

V případě, že je obrazová disparita příliš velká, tak ji mozek není schopen ignorovat a dochází k diplopii. Reakce mozku na tento jev je vergenční a akomodační pohyb. Disparita se tím sníží na přijatelnou úroveň, případně i zmizí. Změnou akomodační vzdálenosti se plátno dostane mimo hloubku ostrosti oka. Jelikož je snímek ale zobrazen pouze na plátně, oči jej vnímají jako rozmazaný. Tak mozek opět začne měnit akomodační vzdálenost, aby byl snímek ostrý. Této změně by měl odpovídat i vergenční pohyb, ale tak se mozek dostane zpět na začátek cyklu. Proto tedy člověk snímek s velkou snímkovou disparitou vnímá ostře a dvojité nebo rozmazaně a jen jednou. Ostré vidění jen jednou je při sledování 3D snímku velmi obtížné, a pokud se pozorovateli podaří, tak dochází ke značné únavě očních svalů. To může mít za následek nevolnost nebo opět dvojité vidění.

Objekty ve 3D snímku tedy mohou vystupovat z roviny snímky pouze v omezené míře. Vystupování objektů je omezeno hloubkou ostrosti oka. Hlavní omezení hloubky 3D snímku je konflikt mezi mechanismy rozpoznávání hloubky, a tedy mezi vergencí a akomodací [3].

## **2.4 Metody stereoskopické projekce**

Stereoskopie je technologie umožňující člověku vnímat prostorový vjem vyvolaný dvourozměrnou předlohou. K vyvolání prostorového vjemu je v první řadě nutné mít vhodná vstupní data, kterými jsou stereo obrazy. Stereo obrazy jsou obrazy (případně videosekvence) zachycující totožnou scénu s určitým horizontálním posunem. Stereo obrazy lze získat snímáním scény pomocí dvou snímačů, které jsou od sebe vzdáleny přibližně jako lidské oči. Pro přenos je možné použít několik formátů. Mimo přímý přenos levého a pravého snímku lze přenášet jeden z nich doplněný hloubkovou mapou, z které lze druhý snímek dopočítat pomocí depth map renderingu [7]. Následně je potřeba zajistit, že každé oko uvidí pouze jemu příslušný obraz (levé oko levý snímek a pravé oko pravý snímek). Pro tento účel je možné využít několik technologií, pomocí kterých je možno vnímat obraz trojrozměrně. Technologie budou stručně popsány v následujícím textu.

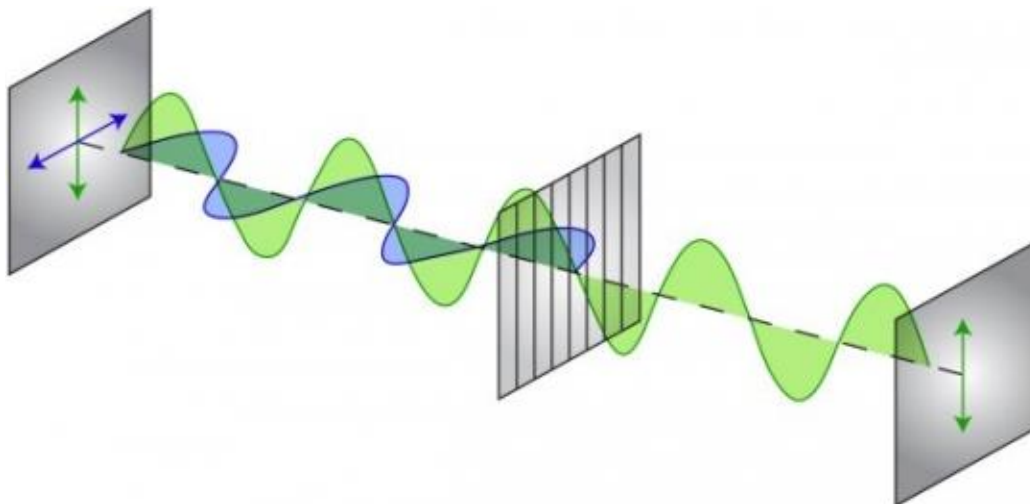
### **2.4.1 Anaglyf**

Tato metoda je nejméně náročná a velmi jednoduchá na provedení, ale zároveň její výsledek je také nejméně kvalitní. Principem této metody je použití speciálních brýlí, kdy očníce mají rozdílnou barvu. Tyto barvy jsou navzájem komplementární (doplňkové). Většinou se používá červená a modrá, popř. zelená. Divák sleduje obraz, který je složený pouze z těchto barev (levý v odstínech jedné z barev a pravý obraz v odstínech druhé barvy) a každé oko tak díky těmto speciálním brýlím vidí pouze jeden obraz. Mozek následně složí tyto obrazy dohromady a divák získá pocit vjemu hloubky obrazu. Nevýhodou této technologie je ztráta informace o barvách.

### **2.4.2 Pasivní 3D projekce**

Mnohem kvalitnější a věrohodnější vjem poskytuje pasivní metoda. Brýle jsou vybaveny polarizačními filtry. Jedna očníce propouští pouze světlo s levotočivou kruhovou polarizací, druhá pouze světlo s pravotočivou kruhovou polarizací (viz.

Obr. 2.1). Obraz je promítán dvěma projektory, přičemž každý z nich vysílá obraz určený pouze pro jedno oko. Před každým z nich je upevněn příslušný polarizační filtr. Každé oko tak vnímá jen obraz, který má a mozek vnímá obraz jako trojrozměrný. Výhodou je možnost použití pro větší počet lidí i relativně nízká cena. Nevýhodou je potřeba dvou projektorů a speciálního nepolarizačního plátna. Na stejném principu může fungovat nejen projektor, ale i TV zobrazovač.



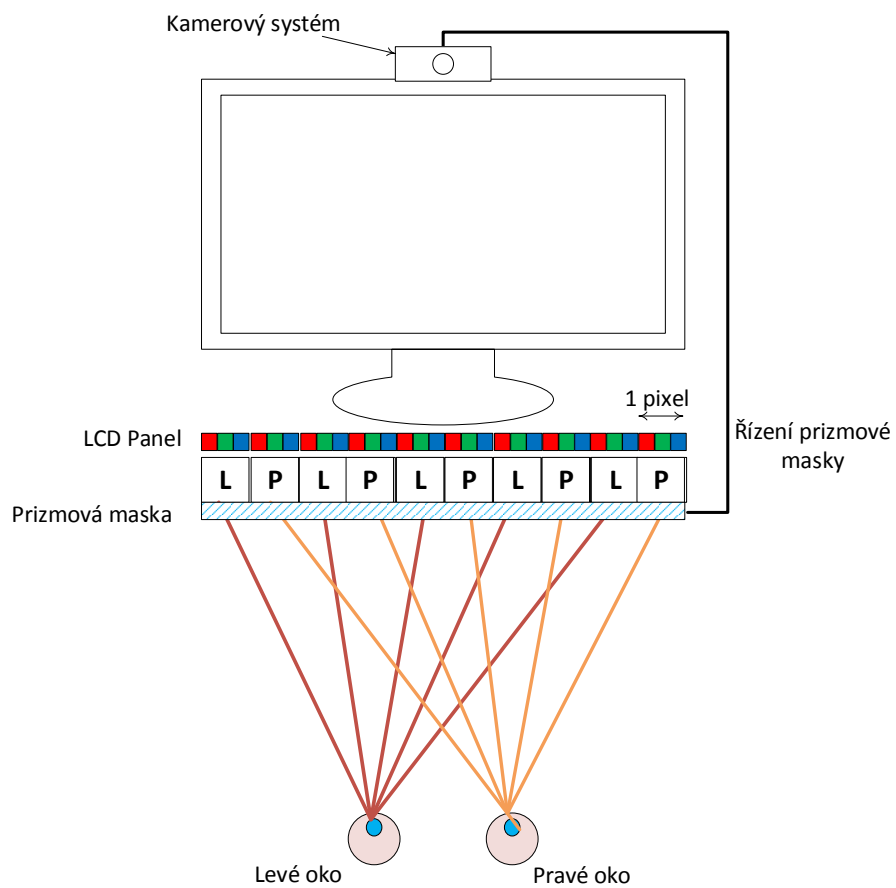
Obr. 2.1: Princip systému s oddělením pomocí polarizace [8].

### 2.4.3 Aktivní 3D systém s časovým multiplexem

Základem této metody je použití dvojnásobné zobrazovací frekvence projektoru či TV zobrazovače, které vysílají (promítají) střídavě obrazy pro levé a pravé oko. Speciální brýle jsou synchronizované se zobrazovacím zařízením a střídavě zatmívají obě očníce ve stejné frekvenci. Díky tomu každé oko tak vidí jen obraz pro něj určený.

### 2.4.4 Autostereoskopické displeje

Tato metoda nevyžaduje použití speciálních brýlí a pro zobrazování se používá LCD displej pokrytý speciální fólií (lentikulární vrstva), která láme směr obrazu. Liché sloupce pixelů jsou lámány jedním směrem, sudé druhým směrem. Každé oko přijme jinou část obrazu a mozek následně vnímá obraz prostorově. 3D vjem je možný pouze v určité poloze diváka vůči displeji [7]. Princip metody je zobrazen na Obr. 2.2.



Obr. 2.2: Princip autostereoskopického zobrazování [8].

### 3 VYMEZENÍ ZÁKLADNÍCH POJMŮ

Tato kapitola je věnována vysvětlení nejdůležitějších pojmů, které se v této práci objeví.

- **Kalibrace.** Termín značí proces, během kterého jsou získány vnitřní a vnější parametry. Vnitřní parametry popisují vlastnosti každé z kamer. Velmi důležitý vnitřní parametr je ohnisková vzdálenost, která může být použita pro přepočítání z disparitní mapy na hloubkovou. Vnější parametry popisují vzájemnou pozici použitých kamer.
- **Korespondující body.** Jsou to dva body v obou obrazech 3D scény reprezentující stejný bod v prostoru.
- **Stereo snímky.** Běžné stereo snímky jsou pořízeny dvěma kamerami majícími rovnoběžné osy snímání (viz. Obr. 3.1a). Pořízené snímky nazýváme levý a pravý snímek. Jejich vzájemná pozice v prostoru se liší v horizontálním směru o určitou vzdálenost, tzv. stereobázi. Pozice korespondujících bodů se liší v horizontálním směru o tzv. horizontální paralaxu (ve vertikálním směru se pozice neliší). Tyto snímky jsou třeba pro další trojrozměrné zpracování.
- **Rektifikace.** Pokud nebyly používány snímky získané z kamer v tzv. normálním ostavení, ale se sbíhajícími se optickými osami či v obecném poloze (viz. Obr. 3.1 b, c) nejsou dostupné stereo snímky, je možné je obdržet procesem rektifikace. Snímky pořízené kamerami v obecné poloze se transformují pomocí nalezení několika korespondujících bodů.
- **Horizontální paralaxa  $p_x$ .** Je to binokulární vodítko, které mozek používá pro vyhodnocení vzdálenosti. Reprezentuje posunutí daného obrazového bodu při srovnání dvou stereo snímků. Jinými slovy je to vzdálenost dvou korespondujících bodů. Větší paralaxu mají bližší předměty, jelikož svou pozici změny více než vzdálenější.
- **Depth Map Rendering.** Cílem této metody je generování obrazu, který je posunutý vůči původnímu v ose  $x$ . Generování probíhá v závislosti na vzdálenosti scény určené z hloubkové mapy. Vytvoří se sekundární virtuální snímek na základě znalosti původního 2D snímku a jemu odpovídající hloubkové mapě a následně se vygeneruje posunutý obraz.
- **Model scény.** Kromě hloubkové mapy lze scénu vyjádřit prostorovou informací i pomocí modelu scény. Popisuje kromě hloubky i zbývající dvě souřadnice daného bodu v prostoru. Model scény je získán pomocí metod 3D rekonstrukce.
- **Stereo disparita.** Jedná se o rozdíl mezi pozicemi korespondujících bodů na stereo snímcích.
- **Disparitní mapa.** Poskytuje informace o hloubce ve scéně. Při znalosti parametrů kamer (ohniskové vzdálenosti  $f$  a stereobáze  $B$ ) je možné vypočítat z disparitní mapy mapu hloubkovou pomocí základního vztahu stereogrammetrie [9]:

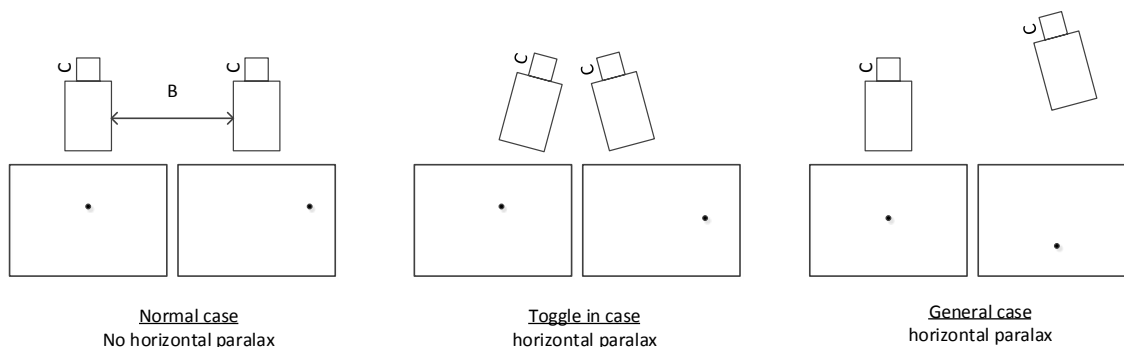


$$Z = \frac{f \cdot B}{d} \quad (3.1)$$

Problém nastává při získávání informací o oblastech, které jsou na krajích snímků, jelikož jsou informace o takové oblasti známy jen z jediného snímku. V takovýchto oblastech vznikají z důsledku nedostatku informace nespojitosti v mapě disparity.

- **Hloubková mapa.** Jedná se o šedotónový obraz reprezentující hloubkové uspořádání scény, a tedy relativní vyjádření hloubky daného bodu korespondujícího s určitým bodem v prostoru. Hloubková mapa je odvozena od horizontální paralaxy daného bodu – čím je paralaxa větší, tím je bod bližší. Platí, že čím má obrazový bod větší jas, tím je blíže, takže světlejší body jsou v popředí a tmavé v pozadí. Nalezení hloubkové mapy pomocí horizontální paralaxy platí pouze ve spojení s pasivní metodou stereo snímání.

Prostorová rekonstrukce úzce souvisí s tvorbou hloubkové mapy, protože v obou procesech je důležitým krokem hledání bodových korespondencí a současně jeden z vnitřních parametrů – ohnisková vzdálenost, ta je použita v přepočtu mezi hloubkovou a disparitní mapou. Z uvedených důvodů je v práci i krátká zmínka o 3D rekonstrukci [8], [9].



Obr. 3.1: Varianty vzájemného umístění kamer v prostoru.

## 4 POUŽITÉ BAREVNÉ MODELÝ A PŘEDZPRACOVÁNÍ

Tvorba (odhad) hloubkové mapy je jednou z metod zpracování obrazu. V průběhu výpočtu hloubkové mapy je často potřeba použít řadu jiných metod pro zpracování obrazu. Výpočet hloubkové mapy svým zaměřením spadá až do poslední fáze zpracování obrazu, a to do pochopení obsahu scény. Proto v této kapitole práce bude stručně popsána samotná reprezentace obrazu a metody použité v dílčích krocích procesu zisku hloubkové mapy. Vzhledem k tomu, že v rámci semestrálního projektu ještě není vytvářena finální aplikace, tak budou stručně popsány ty metody, které budou pravděpodobně v aplikaci použity.

### 4.1 Reprezentace obrazu

Obraz může být reprezentován několika způsoby:

- binární obraz,
- šedotónový obraz,
- barevný obraz.

Každý pixel je v binárním obraze reprezentován jen jedním bitem, a to 0 nebo 1, tudíž barvy v obraze mohou být jen černá a bílá. Používá se především pro definici obrazových masek. Monochromatický šedotónový obraz reprezentuje scénu v odstínech šedi, zpravidla s využitím 8 bitové reprezentace každého bitu hodnotami v rozsahu 0-255. Barevný obraz má několik složek, které závisí na modelu, kterým je tento obraz reprezentován – např. RGB a HSV.

RGB model je využíván převážně v zobrazovacích zařízeních, jako jsou monitory nebo televize. Jeden viditelný bod, tzv. pixel, tvoří tři barevné složky, jež jsou červená, zelená a modrá (Red, Green a Blue). Jejich sloučením vzniknou opticky další barvy. Je to aditivní model barev, tedy pro hodnoty 0,0,0 je výslednou barvou černá. Naopak pro hodnoty 1,1,1 je výsledná barva bílá. Se zvyšujícími se hodnotami v intervalu  $<0;1>$  vznikají další barvy složené ze tří základních složek, které mají různý jas. Pokud jsou všechny tři složky stejné, jedná se o odstín šedi. V počítačové grafice se uvažuje rozsah hodnot každé ze tří složek  $<0;255>$ . Číselné rozmezí pro jednotlivé složky udává tzv. hloubka barev. Osmibitová barevná hloubka je tvořena výběrem 256 definovaných barev. Barevná hloubka o 24 bitech odpovídá počtu  $2^{24}$  barev, nazývá se True Colour a odpovídá přibližně počtu barev, které běžný člověk vidí.

Oproti RGB modelu, který je tvořen základními barvami, HSV model pracuje s vlastnostmi barev, a to s odstínem (tónem), sytostí (saturací) a jasem (Hue, Saturation, Value). Pro zobrazení tohoto modelu se používá šestiboký jehlan umístěný do souřadnicového systému. Na počátku se nachází vrchol jehlanu. Osa jehlanu je shodná se svislou osou znázorňující změnu jasu, který určuje, kolik daná barva odrazí světla, a jak tedy bude jasná. Na vodorovné ose se nachází sytost, která se mění v intervalu  $<0;1>$  a představuje množství přidané bílé složky. Velikost úhlu od vodorovné osy představuje barevný tón. Nevýhodou tohoto modelu je nemožnost plynulého přechodu

mezi černou a bílou nebo plynulé změny barevného tónu. Na druhou stranu model HSV více odpovídá lidskému vnímání barev než ostatní modely. Výhodou je oddělení jasu od barevné informace, což je praktické v mnoha aplikacích v počítačovém vidění [9], [10]. V této práci je využit HSV model právě z důvodu oddělení jasu od barevné informace, což usnadní výpočet disparit.

## 4.2 Použité metody zpracování obrazu

S pořízeným snímkem není možné ihned pracovat, ale musí se nejprve upravit z důvodu zjednodušení další práce s ním. Nejčastější cíle předzpracování obrazu jsou potlačení zkreslení, odstranění šumu, zvýšení kontrastu nebo zdůraznění charakteristik obrazu pro další zpracování [11].

První skupina metod jsou jasové transformace, které se používají převážně ke korekci nerovnoměrného osvětlení, ke korekci kontrastu pomocí ekvalizace histogramu, ke zvýšení dynamického rozsahu nebo gama korekci. Vstupem i výstupem je obraz, který má stejné parametry. Změní se hodnoty obrazové funkce vstupního obrazu podle daného pravidla. Jasové transformace lze rozlišit dle velikosti okolí vyšetřovaného bodu na globální, lokální a pixelové [12].

Další z metod je tzv. segmentace, což je rozčlenění obrazu do částí, které souvisí s předměty či oblastmi reálného světa. Prakticky je to oddělení objektů od pozadí a následná analýza obsahu obrazu. Vstupem je obraz, výstup může být různý dle použité metody (např. část obrazu, poloha objektu v obraze či obraz). Segmentací chystáme obraz na další krok, jímž je popis.

Detekce hran může být součástí procesu segmentace. Je to metoda, při které se zdůrazní vyšší frekvence v obraze, a tím se zvýrazní obrazové elementy, kde se jasové funkce náhle mění – zvýrazní se tedy hrany. Hrana je kolmice gradientu, což je směr největšího růstu obrazové funkce od černé po bílou. Je určena velikostí gradientu obrazové funkce a směrem [13]:

$$|\nabla g| = \sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2}, \quad (4.1)$$

$$\varphi = \arctg \left( \frac{\frac{\partial g}{\partial y}}{\frac{\partial g}{\partial x}} \right). \quad (4.2)$$

## 5 METODY VÝPOČTU DISPARITNÍ MAPY

Jelikož se pro tvorbu hloubkových map používají stereo snímky, popř. rektifikované snímky, tak se prostor změní z trojrozměrného na dvourozměrný. Vzhledem k tomu, že korespondující body musí být na stejném řádku, tak se celý vyhledávací prostor zmenší jen na vyhledávání v jednorozměrném prostoru.

Jeden snímek páru je vždy považován za referenční, tzv. „base“ a druhý za porovnávaný, tzv. „match“. Podle toho se rozlišují metody left-to-right nebo right-to-left stereo porovnávání.

Cílem algoritmu je tvorba hloubkové (disparitní) mapy, tedy mapy odhadující disparitu pro každý pixel. Odpovídající si pixely mají nulovou disparitu. Porovnávání jediného pixelu by nevedlo k dobrým výsledkům, protože by se na sousedních bodech objevily velmi rozdílné hloubky. Je tedy nutné využít složitějších postupů, které porovnávají celé skupiny pixelů. Předpokladem je, že v obraze budou velké plochy s podobnou disparitou a jen málo výrazných rozdílů, kdy se mění výrazně hloubka.

Pixelovým souřadnicím  $(x, y)$  referenčního snímku odpovídají souřadnice  $(x', y')$  na snímku. Pro jejich vzájemný vztah platí [14]:

$$x' = x + s \cdot d(x, y), y' = y, \quad (5.1)$$

kde  $s = \pm 1$  je znaménko zvolené dle konvence tak, aby disparita byla vždy kladná kvůli snazším výpočtům,  $d(x, y)$  je disparita každého bodu. Na podobném principu pracuje většina algoritmů, jen pro určité fáze výpočtu používá funkce s obdobným významem. Tyto algoritmy fungují tak, že pro několik disparit určí míru podobnosti nebo nepodobnosti intenzit korespondujících bodů. Počítají několik obrazů – pro každý je definováno konstantní posunutí. Pokud jsou body shodné, disparita je nulová. Hodnoty míry podobnosti všech pixelů a všech disparit tvoří prostor disparitních obrazů, který se označuje  $C(x, y, d)$ . Cílem algoritmu je nalézt co nejoptimálnější výběr disparit tak, aby odpovídal co nejvíce skutečnosti. Úlohou tedy je vybrat z obrazů ten nejlepší.

Právě způsob určení míry podobnosti, vyrovnání se s rychlými změnami podobnosti mezi sousedními pixely a výběru nejoptimálnějšího výběru disparit je pro každý z algoritmů specifický.

Tvorbu disparitní mapy lze obecně rozdělit do následujících pěti fází

- 1. Předzpracování**
- 2. Výpočet mír podobnosti pro veškeré dané disparity neboli náklad či funkce porovnání, tzv. „matching cost“.**
- 3. Agregace (seskupení) vypočtených mír při dané disparitě, tzv. „cost aggregation“.**
- 4. Určení výsledné nejoptimálnější disparity pro každý pixel.** Existují dvě skupiny metod, které lze v tomto kroku využít – lokální a globální. Lokální metody hledají disparity jen v určitém okolí daného bodu. Globální metody hledají vhodný výběr disparit v celém obraze, jež je považován za jeden celek. Podrobněji jsou lokální a globální metody popsány v dalších kapitolách.

## 5. Doladění výsledné hloubkové mapy [14], [15].

V další části se práce bude podrobněji zabývat jen nejdůležitějšími kroky, tedy výpočtem mír podobnosti, agregací a určením výsledné neoptimálnější disparity. Kroku předzpracování byla věnována část kapitoly 4.

Existuje bezpočet algoritmů pro výpočet hloubkových map. I řešení, která jsou založena na stejném principu, lze naprogramovat mnoho způsoby. Další kapitoly se zabývají převážně metodami, jež budou v praktické aplikaci použity nebo jsou významné.

### 5.1 Výpočet mír podobnosti

Porovnávají se vlastnosti jednotlivých pixelů, např. jejich jas. Existují různé metody porovnávání založené na různých matematických přístupech. Pro výpočet se používají metody jako Squared Intensity Differences, což je kvadrát rozdílů jasu, Sum of Squared Differences – suma rozdílů čtverců, Absolute Intensity Differences – absolutní hodnota rozdílů intenzit, Mutual Information – vzájemná informace nebo Structural Similarity Index vyjadřující podobnost dvou obrazů. Další úpravou je saturace, při níž hodnoty, které jsou vyšší než předem definovaná hodnota, se sníží na tuto hodnotu. Tato úprava je prováděna z důvodu lepšího výsledku v druhém kroku. Dalšími metodami výpočtu míry podobnosti jsou normovaná korelace či binární porovnání. Algoritmus se snaží najít větší oblasti s celkově nejvyšší mírou podobnosti. Taková metoda je mnohem úspěšnější než ta, jež porovnává samostatné pixely.

- **Absolutní rozdíl (AD, z angl. absolute difference)**

Jedná se o absolutní rozdíl jasových hodnot vypočtený dle vztahu [16]

$$AD = |I_L(x, y) - I_R(x + d, y)|, \quad (5.2)$$

kde  $I_L(x, y)$  je jasová hodnota pixelu o souřadnicích  $x, y$  v referenčním snímku a  $I_R(x + d, y)$  je jasová hodnota pixelu o souřadnicích  $x + d, y$  v porovnávaném snímku. Pozice v porovnávaném snímku je postupně měněna vlivem velikosti aktuálně testované disparity.

- **Rozdíl čtverců**

$$AD = I_L(x, y)^2 - I_R(x + d, y)^2. \quad (5.3)$$

Význam proměnných je shodný jako ve vztahu 5.2 [16], [17].

- **Vzájemná informace (MI, z angl. Mutual Information)**

Jedná se o míru závislosti proměnných. Tato metoda se používá například při korekci rozdílů mezi obrazy. Maximalizace vzájemné informace vede k minimalizaci funkce zjištění nákladů. Definice této metody dle vztahu [18]

$$E_{data}(f) = -MI(I_L, I_R, f). \quad (5.4)$$

Význam proměnných je shodný jako ve vztahu 5.2 [18].

- **SSIM – Podobnost dvou obrazů (Structural Similarity Index)**

SSIM je index podobnosti obrazů, který nabývá hodnot -1 až 1, kde 1 vyjadřuje úplnou shodu obrazů. Je založen na výpočtu tří členů – jasu, kontrastu a struktuře. Výsledný index je multiplikace všech těchto členů [19]:

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma \quad (5.5)$$

Kde  $l(x, y)$  je jasová hodnota pixelu o souřadnicích  $x, y$  v referenčním snímku,  $c(x, y)$  je hodnota kontrastu pixelu a  $s(x, y)$  je informace o struktuře pixelu. Ale obvykle se index počítá jen vůči jasové složce [19].

- **CGRAD**

Metrika používá stejný výpočet jako metrika absolutních diferencí s tím rozdílem, že se difference nepočítají v jasovém obraze, ale v gradientním, který je získán z jasového výpočtem gradientů ve všech směrech [20]:

$$C_{GRAD}(x, y, d) = \sum_{(i,j) \in N_x(x,y)} |\nabla_x I_1(i, j) - \nabla_x I_2(i + d, j)| + \sum_{(i,j) \in N_y(x,y)} |\nabla_y I_1(i, j) - \nabla_y I_2(i + d, j)| \quad (5.6)$$

Kde  $N(x, y)$  je obklopující okno ve velikosti  $3 \times 3$  na pozici  $(x, y)$ ,  $N_x(x, y)$  je obklopující okno bez sloupce umístěného úplně vpravo,  $N_y(x, y)$  je obklopující okno bez nejspodnější řady,  $\nabla_x$  je gradient v pravém směru,  $\nabla_y$  je gradient ve směru dolů. Pro barevné obrazy je výpočet stejný, jen se sečtou rozdílnosti měření pro všechny kanály [20].

## 5.2 Agregace vypočtených mír

Používá se téměř výlučně jen u lokálních metod. Úkolem agregace je odstranit drobné výkyvy. Nesrovnávají se podobnosti jednotlivých pixelů, ale podobnosti oken složených z více pixelů, které jsou typicky čtvercové o známém rozměru. Jedná se následně většinou o součet nebo průměr dané oblasti. Typickou metodou je Sum of absolute differences, což je suma absolutní rozdílů. Podrobněji je rozvedena v následující kapitole. Metody, které využívají informace ze všech disparitních obrazů, používají omezení změny disparity nebo omezení gradientu disparity.

U pixelů s podobnými vlastnostmi je možné je snadno zaměnit díky zkreslení hodnot například vlivem šumu. To vede k chybnému vyhodnocení a chybovému pixelu v hloubkové mapě. U větších ploch je tento problém méně znatelný. Rozdíl porovnávání oken oproti porovnávání jednotlivých pixelů spočívá v zanedbání stavu v rámci jednotlivých oken, jelikož jsou pro výpočet důležité jen jeho celkové vlastnosti a vnitřní stavbu okna je možno považovat za homogenní. Důsledkem je ztráta detailů a menší podrobnost výsledné hloubkové mapy. Problém nastává také při vyhodnocování ploch s konstantními vlastnostmi nebo s opakujícími se motivy. V takových případech lokální metody nedokáží vybrat s jistotou odpovídající si okna, jelikož jich existuje v dané oblasti se shodnými vlastnostmi více.

Mezi další problémy, které nastávají u algoritmů s porovnávacími okny, patří vyhodnocování přítomnosti objektů, které jsou ohnuté či zkosené. V takovém případě by bylo vhodné měnit tvar porovnávacího okna. Chybné vyhodnocení nastává také,

když porovnávací okno překrývá plochu s měnící se hloubkou, jelikož algoritmy uvažují hloubku v rámci okna konstantní. Problém je i výskyt motivu, který je na druhém snímku zakrytý. Tento motiv pak nelze nikde nalézt a výsledek je tak chybný. Vyjmenované problémy lze částečně vyřešit rozdělením porovnávacího okna do více sektorů nebo vytvořením posunutelného centra okna. Lokální metody jsou rychlejší a méně výpočetně náročné, ale jejich výsledky bývají v porovnání s globálními horší [15], [17].

## 5.3 Určení výsledné neoptimálnější disparity

Existují dvě skupiny metod, které lze v tomto kroku využít – lokální a globální. Lokální metody hledají disparity jen v určitém okolí daného bodu. Globální metody hledají vhodný výběr disparit na celém řádku, jež je považován za jeden celek.

### 5.3.1 Lokální metody

Lokální metody jsou výpočetně jednodušší – výsledná disparita se vybírá způsobem tzv. „winner-take-all“, což znamená, že se hledá nejlepší výsledek, který se použije, zatímco se ostatní výsledky neberou v potaz. Vybraná nejlepší disparita pro každý pixel je ta, jež má nejmenší agregovanou míru podobnosti. Agregace zajišťuje souvislejší plochy a vylučuje jednopixelové chyby [13].

### 5.3.2 Globální metody

První kroky, tedy předzpracování a určení hodnot funkce porovnání, se využívají stejně jako u lokálních metod. Ale další krok – seskupení hodnot funkce porovnání, se týká už jen metod lokálních. Globální metody neuvažují vlastnosti pouze jednoho pixelu, ale i okolí bodu a potažmo celého obrázku. Navíc také zohledňují fakt, že v blízkém okolí se hloubka skokově příliš nemění a jsou tak méně náchylné k chybám. Je možno je rozlišit na jednorozměrné, které zkoumají pouze vlastnosti pixelů v rámci jedné přímky, což je například dynamické programování, nebo dvourozměrné – zohledňují i vlastnosti obrazu dané jeho druhým rozměrem – např. řezy grafů. Tyto algoritmy jsou komplexnější a výpočetně náročnější, ale výsledky produkují lepší.

Globální metody jsou zaměřeny převážně na minimalizaci energie. Zohledňují při hledání shod kromě vlastností daného pixelu i to, jaké shody byly přiřazeny k okolním pixelům. Tedy přiřazení korespondujících bodů k sobě ovlivňuje přiřazování ostatních bodů v jejich okolí. Jejich úkolem je najít disparitní funkci  $d$ , která minimalizuje globální energii [15]:

$$E(d) = E_{data}(d) + \lambda \cdot E_{smooth}(d). \quad (5.7)$$

Data term  $E_{data}(d)$  měří míru shody disparitní funkce  $d$  s vstupním obrazovým párem [15].

$$E_{data}(d) = \sum_{(x,y)} C(x, y, d(x, y)). \quad (5.8)$$

Prostor disparitních obrazů může být dle situace počáteční nebo agregovaný.

Druhý term se téměř vždy omezuje jen na měření rozdílů disparit mezi sousedními pixely [15].

$$E_{smooth(d)} = \sum_{(x,y)} \rho(d(x,y) - d(x+1,y)) + \rho(d(x,y) - d(x,y+1)), \quad (5.9)$$

kde  $\rho$  je monotónně vzrůstající funkce rozdílů disparit.

Jakmile je energie definovaná, je možno aplikovat algoritmy pro její minimalizaci. Patří mezi ně metody např. simulované ochlazování, max-flow nebo graph-cut.

Dalším přístupem ke globální optimalizaci je tzv. dynamické programování. Hledá minimální cenu cesty v matici tvořené korespondujícími řádky. Možností je také využít tzv. kooperativní algoritmy, které iterativně provádí lokální výpočty vedoucí ke globální minimalizaci [15].

- **Scanline algorithm**

Scanline algoritmy jsou zvláštním druhem geometrické transformace a pracují pouze se sloupci a řádky. Po převzorkování skrz Scanline algoritmus se z problému stává jednorozměrný problém, což představuje velké zjednodušení. Program operuje vždy jen s jednou dimenzí. Od ostatních globálních algoritmů se Scanline liší tím, že se zaměřuje na vztah minimalizace energie a lokální shody funkce nákladů porovnání. Lokální řešení je transformováno do minimalizace energie přes optimalizaci disparitních nespojitostí. Scanline algoritmus je možné rozdělit do dvou kroků. Nejprve je úvodní disparitní mapa obdržena bez použití nespojité funkce porovnání. Následně je disparitní mapa optimalizována iterativně podle funkce energie.

Energie je formulována dle tohoto vzorce [21]:

$$E(f) = \sum_{(p,q) \in N} V(f_p, f_q) + \sum_{p \in P} D_p(f_p). \quad (5.10)$$

Okolí  $N$  obsahuje jen sousední pixely z levé a pravé strany.  $V(f_p, f_q)$  je disparitní nespojitost mezi body  $p$  a  $q$ .  $D_p(f_p)$  je funkce porovnání označená  $f$  při bodu  $p$ .

Komplikace minimalizace energie nastává kvůli ceně nespojitosti. Pro příklad je použita tato funkce [21]:

$$E_d(f) = \sum_{p \in P} D_p(f_p). \quad (5.11)$$

Minimum této funkce je triviální, jelikož lokální minimum je zde také globálním minimem. Je jasné, že minimalizace vztahu 4.7 bude mít větší disparitní nespojitost než 4.6. Minimální řešení 4.7 může být přetransformováno do minimálního řešení 4.6 eliminací a relokací nespojitostí disparit. Pro jednodimenzionální obraz se nalezení optimální pozice nespojitostí stává triviálním úkolem [21].

- **Belief Propagation**

Belief propagation je algoritmus, který byl vyvinut především s cílem najít mezní pravděpodobností v bayesovských sítích. Algoritmus této metody může být popsán schématem na obr. 5.1. Základním principem je šíření informace přes grafický model pomocí zasílání zpráv mezi sousedními uzly. Zasílaná zpráva je iterativně upravována. V případě zpracování obrazu uzel reprezentuje pixel obrazu a každý z uzlů má čtyři sousedy. V každém iteračním kroku každý pixel spočítá informaci na základě zprávy



doručené od svých sousedů a získanou informaci pošle všem svým sousedům. V následujícím odstavci je stručně popsáno využití tohoto algoritmu pro tvorbu hloubkové mapy [22].

Postup tvorby hloubkové mapy pomocí belief propagation si lze představit jako postupné štítkování (labeling) pixelů v obraze. Postup lze nejlépe vysvětlit na příkladu. Jako vstup máme levý a pravý snímek scény a každý ze snímků má  $M$  řádků a  $N$  sloupců. Dále existuje sada štítků, které představují hodnoty disparity  $L = \{0, 1, 2, \dots, d_{\max}\}$ , kde  $d_{\max}$  je největší možná disparita vyskytující se v obraze. Naším cílem je označit všechny pixely levého obrazu jedním ze štítků ze sady  $L$ . Za tímto účelem použijeme následující základní rovnici belief propagation [23]

$$b_q(f_q) = D_q(f_q) + \sum_{p \in N(q)} m_{p \rightarrow q}^T(f_q), \quad (5.12)$$

kde  $p$  a  $q$  představují sousední pixely,  $N(q)$  označuje okolí pixelu  $q$  a  $f_q$  je hodnota disparity.  $D(f_q)$  je cena přiřazení disparity  $f_q$  pixelu  $q$ . Tato cena přiřazení může být definována (spočítána) jako rozdíl intenzit korespondujících pixelů levého a pravého snímku. Rovnice pro výpočet  $D(f_q)$  je následující [23]

$$D_q(f_q) = I_{\text{LEFT}} - I_{\text{LEFT}+f_q}, \quad (5.13)$$

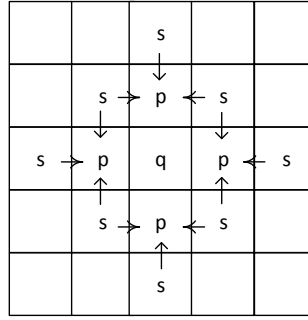
kde  $I_{\text{LEFT}}$  je hodnota intenzity pixelu v levém obraze a  $I_{\text{LEFT}+f_q}$  je hodnota intenzity bodu v pravém obraze (pozice bodu je posunuta o hodnotu testované disparity).

Předpokládejme, že ve vzorovém příkladu se vyskytnou čtyři různé hodnoty disparity označené štítky  $L_{d0}, L_{d1}, L_{d2}, L_{d3}$ . Pro každý štítek (hodnotu disparity) je vytvořena informační matice pro šíření zprávy. Každá pozice v matici reprezentuje jeden pixel obrazu. V prvotním okamžiku nastavíme hodnotu  $A = I_{L1} - I_{R1}$  do informační matice štítku  $L_{d0}$ ,  $B = I_{L1} - I_{R2}$  do informační matice štítku  $L_{d1}$ ,  $C = I_{L1} - I_{R3}$  do informační matice štítku  $L_{d2}$  a  $D = I_{L1} - I_{R4}$  do informační matice štítku  $L_{d3}$ . Následuje iterační krok. Během každé iterace jsou informační matice změněny pomocí následující rovnice [23]

$$\sum_{p \in N(q)} m_{p \rightarrow q}^T(f_q) = \min_{f_p} \left( v(f_p - f_q) + D_p(f_p) + \sum_{s \in N(p)} m_{s \rightarrow p}^{t-1}(f_p) \right). \quad (5.14)$$

Rovnice počítá zprávu posílanou z pixelu  $p$  do pixelu  $q$  v iteračním kroku  $t$ . Na pravé straně rovnice je součet ceny přiřazení disparity  $f_p$  pixelu  $p$  a rozdílu disparit  $f_p$  a  $f_q$ . Poslední část rovnice je součet zpráv přicházejících do pixelu  $p$  v předchozí iteraci.

Iterační krok je opakován, dokud není splněno kritérium. Když je iterační fáze ukončena, tak dostáváme nové hodnoty v každé buňce informačních matic  $A$ ,  $B$ ,  $C$  a  $D$ . Pro závěrečné přiřazení správné disparity (štítku) každému pixelu vybereme disparitu, pro niž je dosaženo nejmenší ceny přiřazení štítku [23].



Obr. 5.1: Základní princip šíření zprávy v belief propagation [23].

### • Mean Shift Algorithm

Tento algoritmus shlukuje pixely na základě jejich podobnosti a blízkosti pozice pomocí konvergence do lokálních maxim spojeného intenzitního a souřadnicového prostoru. Pixely, o nichž se předpokládá, že patří jednomu objektu, tvoří tzv. shluky. Vzdálené jsou naopak pixely, které mají různou intenzitu a patří tak pravděpodobně k různým objektům. V každém kroku se tedy algoritmus posune ve směru váženého průměru vzorků z oblasti současné pozice.

Úloha segmentace se zde tedy převede na shlukování bodů.

Pro iterativní nalezení lokálního maxima hustoty pixelů se používá obecný algoritmus Mean Shift. Má dvě části:

1. Mean Shift je spuštěn u každého vzorku a zapamatuje si lokální maximum, do kterého konvertoval.
2. Shlukují se ty vzorky, které dokonvergovaly přibližně do stejného maxima. Oblast, ze které vzorky konvergují do stejného bodu, se nazývá „basin of attraction“.

Algoritmus v každém kroku počítá gradient hustoty vzorků v současné pozici  $x$  ze vzorků v okolí, které je definováno tzv. kernelem  $K$ , což je radiálně symetrická funkce. Pro obecný kernel je odhad hustoty v bodě  $x$  [15]

$$f(x) = \frac{1}{n \cdot h^d} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right), \quad (5.15)$$

kde  $h$  je šířka kernelu a  $x_n$  jsou vzorky. Jelikož je ale potřeba jen gradient, který se získá [15]

$$\nabla f(x) = \frac{2 \cdot c_k}{n \cdot h^d} \sum_{i=1}^n (x - x_i) \cdot k' \left( \left\| \frac{x - x_i}{h} \right\|^2 \right). \quad (5.16)$$

Mean shift algoritmus, který byl inicializován v pozici  $y_0$ , proběhne tímto způsobem [15]:

$$\begin{aligned} &do \\ &y_{i+1} \leftarrow y_i + \nabla f(y_i) \\ &i \leftarrow i + 1 \\ &until \|y_{i+1} - y_i\| < \frac{h}{1000} \end{aligned} \quad (5.17)$$

Pokud jsou použity různé veličiny na různých osách, je vhodné použít kernely s různou šířkou pro každou veličinu. Pro segmentaci obrazu je vhodná možnost nezávislé změny kernelu ve směru pozice pixelu a kernelu ve směru intenzity. Pro každý pixel dostaneme lokální maximum  $z=[u, v, i]$ . Je vhodné spojit takové sousední pixely  $a, b$ , pro něž platí ve všech souřadnicích [15]

$$abs(z_a - z_b) \leq h = [h_s, h_s, h_r]. \quad (5.18)$$

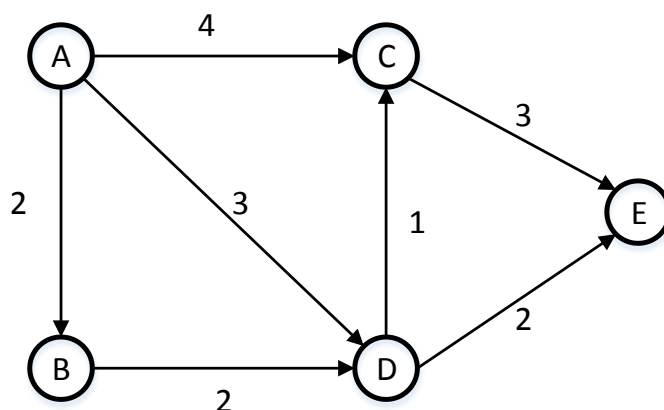
Ke spojení se používá technika tzv. supergrid, kdy se vytvoří matice o dvojnásobné délce stran, než má obraz a jsou v ní reprezentovány pixely i hrany. Pokud je splněna podmínka 5.18, tak hrana nabývá hodnoty 1. V opačném případě je hodnota 0.

Algoritmus Mean Shift je časově náročný a doporučuje se tedy pracovat během vývoje se zmenšeným obrazem [15].

- **Graph Cut**

Jedná se o stereo porovnání s globálním přístupem. Při hledání korespondujících bodů tato metoda nehledá vlastnosti bodů pouze po přímce, ale řeší i vlastnosti okolí. Při předpokladu, že se hloubka, a tedy i hodnoty disparity, nemění skokově a sousední pixely nabývají podobných hodnot, přiřazování korespondujících bodů ovlivňuje i přiřazování bodů sousedních.

Nejprve se vybere referenční snímek ze dvou možných stereo snímků. Následně je nutná definice funkce nákladů porovnání („matching cost function“). Po těchto dvou krocích je možno přistoupit k samotnému hledání hloubkové mapy. Řešený problém se popíše graficky pomocí trojrozměrné mřížky. Od ní lze pomocí vhodné transformace přejít ke specifickým grafům, které se skládají z uzlů odpovídajících jednotlivým pixelům a ze dvou speciálních přidáných uzlů source (S) a sink (T). Jedná se tedy o síť propojených uzlů. Síť se následně rozdělí na dvě části, a to podle hodnot disparity, kterých nabývají jednotlivé pixely. Hledá se tzv. optimální řez. K uzlu source (S) pak zůstanou připojeny uzly, které odpovídají pixelům s vyšší hodnotou disparit [22].



Obr. 5.2: Schéma algoritmu graphcut [22].

- **Semiglobální metoda využívající vzájemnou informaci**

I další kapitola popisuje konkrétní metodu navrženou v článku [24]. Metoda využívá semi globální hledání korespondencí a pro porovnání podobnosti používá pixelově orientovanou metriku vzájemné informace (MI) popsanou v kapitole 5.1. Metrika MI

by měla být schopná kompenzovat radiometrické rozdíly mezi oběma vstupními snímky.

Pixelově orientované hledání shody je doplněno vyhlazovací funkcí, což je obvykle označováno jako globální funkce. Metoda provádí rychlé aproximace pomocí všesměrových optimalizací. Metoda se rovněž zabývá řešením detekce okluzí či subpixelovým zpřesněním. Kromě toho, metoda provádí na závěr kroky pro odstranění odlehlých (nepravděpodobných) hodnot a interpolaci neurčených pixelů. Autoři navrhuji strategie pro zpracování téměř libovolně velkých obrazů a fúze disparitních snímků pomocí ortografické projekce. Složitost roste lineárně s počtem pixelů a disparitním rozsahem, což znamená rychlé výpočty.

## 5.4 Konkrétní algoritmy

### 5.4.1 Depth Discontinuities by Pixel-to-Pixel Stereo

Nejedná se o klasickou aplikaci subpixelových výpočtů, tedy pixel jednoho obrazu  $x_i$  není porovnáván vůči pixelu  $y_i$  v obraze druhém. Provede se interpolace a zjišťuje se minimální a maximální hodnota intenzity v oblasti o půl pixelu na každou stranu. Pokud je zkoumaná hodnota intenzity v zjištěném intervalu, je podobnost dokonalá. Pokud ne, je spočtena míra podobnosti. Takto je zajištěno lepší párování a rozdílnost zesílení kamer nemá takový vliv na kvalitu.

Určení finální disparity je počítáno za pomoci dynamického programování. Funkce popisující energii [25]:

$$\gamma(M) = N_{OCC} \cdot \kappa_{OCC} - N_m \kappa_r + \sum_{i=1}^{N_m} d(x_i, y_i) \quad (5.19)$$

, kde  $N_{OCC}$  a  $N_m$  jsou počty spárovaných a nespárovaných úseku (ne pixelů)

$\kappa_{OCC}$  a  $\kappa_r$  jsou konstantní penalty nespárovaných a odměny spárovaných úseků

Dále jsou pro tento algoritmus použity další doladující úpravy. Například podle disparity se očekává, že nespárované pixely se objeví v jednom z obrazů podle toho, kde je objekt bližší. Také během dynamického programování je prohledávaný prostor zmenšován a jednotlivé cesty optimalizovány [25].

### 5.4.2 SSD

I přesto, že se jedná se velmi jednoduchý algoritmus, poskytuje kvalitou přijatelné výsledky. Dle rozdělení fází v předchozí kapitole jde tento postup shrnout do tří bodů:

1. Míra podobnosti je kvadrát rozdílů intenzit při dané disparitě
2. Agregace je součtem mír přes čtvercové okno při dané disparitě
3. Výsledná disparita je vybrána jako minimální agregovaná hodnota pro každý pixel (Winner-Take-All) [26]

## 5.5 Využití korespondencí výrazných bodů obrazu

Drtivá většina výše zmíněných metod vychází z hledání bodových korespondencí pro každý pixel obrazu bod po bodu. Nalezení korespondujícího bodu pro pixel, jenž se neodlišuje od okolních pixelů, může být velmi obtížné a v důsledku nespolehlivé. Odlišným přístupem může být hledání bodových korespondencí pouze pro výrazné body obrazu. Metody pracující na tomto principu se v anglické literatuře nazývají „feature based“ [27]. Výrazné body obrazu jsou takové body, které se výrazně liší od svého okolí a jsou tak dobře identifikovatelné [28]. Výrazný obrazový bod musí být nalezitelný opakovaně a spolehlivě. Z toho vyplývá, že daný bod můžeme lehce najít v obou stereo obrazech a tím nalézt korespondující body a určit disparitu daného bodu. K hledání výrazných bodů obrazu (scény) jsou používány tzv. detektory a deskriptory. Postup lze rozdělit do následujících kroků

- Nalezení sad výrazných bodů v obou obrazech
- Výpočet deskriptoru výrazných bodů
- Porovnání deskriptorů jednotlivých bodů- určení korespondencí

Mezi nejznámější a nejčastěji používané detektory patří SIFT a SURF [29]. Algoritmy by měli být odolné vůči natočení scény, změně měřítku, změně jasové hodnoty.

Takto nalezené korespondence však nebudou souvisle pokrývat celý obraz (scénu) a nevznikne souvislá hloubková mapa. Z toho důvodu musí být provedeny další kroky pro získání souvislé hloubkové mapy (v angl. literatuře dense depth map). Logickým navazujícím krokem je segmentace scény a následné využití segmentovaného obrazu takovým způsobem, že disparita konkrétního bodu scény (výrazného bodu) je přiřazena celému segmentu, jehož je bod součástí. Získané disparity výrazných bodů lze popřípadě využít k zpřesnění nebo doplnění hloubkové mapy získané s využitím hledání bod po bodu [29].

## 6 IMPLEMENTACE METOD

Po získání teoretických znalostí byly vybrány metody reprezentující lokální a globální metody pro tvorbu hloubkových map, tyto metody byly implementovány v prostředí MATLAB.

### 6.1 Lokální metoda

Pro implementaci lokální metody byl použit základní algoritmus pro tvorbu hloubkové mapy tak, jak je popsán v kapitole 5.1. Algoritmus může být popsán vývojovým diagramem na obrázku Obr. 6.1. Základ algoritmu je převzatý z implementace volně ke stažení [31]. Kód byl mírně pozměněn a byla přidána další metrika. Základními krokem je hledání správné hodnoty disparity pomocí míry podobnosti a agregace této míry podobnosti v různě velkém okně. V prvním kroku jsou definovány parametry algoritmu  $P_{weigh}$ , jenž určuje váhu míry podobnosti vypočítané z gradientních snímků (podrobněji vysvětleno v dalším textu),  $win\_size$  definující velikost okna pro agregaci míry podobnosti (také možné pojmenovat jako velikost průměrujícího okna) a  $d_{max}$  definující maximální velikost disparity. V dalších dvou krocích je opakovaně dvakrát volána funkce *SlideImages*

```
[dsp1, diff1, DSI1] = slide_images(i1,i2, 1, maxs, win_size, weight),
```

kteřá je jádrem algoritmu, v rámci něhož se provádí evaluace míry podobnosti i agregace podobnosti v určitém okolí definovaném parametrem  $win\_size$ . Vstupy funkce jsou levý a pravý snímek scény a výše popsané parametry lokální metody. Výstupy funkce jsou disparitní mapa  $dsp$ , matice minimálních cen disparit pro každý pixel  $diff$  a trojrozměrná matice  $DSI$ , která uchovává ceny všech uvažovaných disparit pro každý obrazový bod. Proměnná  $DSI$  je využívána jako vstup pokročilejších optimalizačních metod popsanych v dalších kapitolách. Funkce je volána dvakrát, protože jednou jsou vyhledávány bodové korespondence ve snímku  $i_2$  pro pixelu ve snímku  $i_1$  a podruhé je tomu naopak. V důsledku jsou tedy získány dvě disparitní mapy  $dsp_1$  a  $dsp_2$ . V předposledním kroku postupu je volána funkce *WinnerTakeAll*

```
[dsp] = winner_take_all(dsp1,diff1,dsp2,diff2);
```

Vstupem funkce jsou dílčí disparitní mapy  $dsp_1$  a  $dsp_2$  a matice cen zvolených disparit  $diff1$  a  $diff2$  získané v předešlém kroku. Výstupem je finální disparitní mapa. Úloha této funkce je přiřadit každému pixelu tu nevhodnější hloubku, tedy buď hodnotu mapy  $dsp_1$  nebo  $dsp_2$ . Rozhodnutí je provedeno na základě matic  $diff1$  a  $diff2$ , které obsahují cenu dané disparity v jedné a v druhé disparitní mapě. Jako finální hodnota disparity pro daný pixel je zvolena hodnota z té disparitní mapy, ve které je cena disparity pro daný pixel menší. V posledním kroku je provedena snaha eliminovat alespoň nějaké chyby, které spočívají v nespojitosti hloubkové mapy. Tohoto cíle je dosaženo pomocí průměrování hodnot v okolí daného bodu.

Stěžejní funkce *SlideImages* je popsána samostatným vývojovým diagramem na obrázku 6.2. V prvním kroku jsou vytvořeny výstupní matice  $dsp$  obsahující zjištěnou nejvýhodnější disparitu a  $diff$  obsahující cenu (hodnotu míry podobnosti) dané disparity,

která je v průběhu algoritmu postupně aktualizována cenou průběžně vítězné disparity. Matice disparit *disp* je na počátku naplněna samými nulami. Matice cen disparit je naplněna hodnotami *inf* (tedy nekonečně vysokými hodnotami), a to z toho důvodu, že jako vítězná disparita pro daný pixel je zvolena ta disparita, jenž má nejnižší cenu, a tudíž potřebujeme, aby při prvním porovnání stávající a nové hodnoty ta nová byla menší, nahradila novou a stanovila tak skutečnou počáteční cenu disparity.

V následujícím kroku jsou vypočteny gradientní obrazy obou snímků, a to ve všech směrech: horizontální (v řádku), vertikální (ve sloupci) a napříč barevnými složkami

```
[glx gly glz] = gradient(double(i1))
```

V následujícím kroku je snímek, v němž je korespondence hledána, posunut celý o hodnotu disparity *d* pomocí funkce *shift\_image*, která naplní okrajové dolní sloupce

```
s = shift_image(i2,i),
```

kde *i* udává velikost posunu snímku a je daný testovanou disparitou a *i2* je vstupní barevný snímek. Disparity jsou postupně procházeny pomocí funkce *for* od 0 po  $d_{max}$ . V dalším kroku je vypočtena míra podobnosti v barevném prostoru RGB pro všechny pixely obrazu při dané disparitě (mírou podobnosti je buď AD nebo SSIM)

```
diffs = sum(abs(i1-s),3) %výpočet podobnostimetrikou AD.
```

V dalším kroku je vypočtena míra podobnosti v gradientních obrazech,

```
gdifffx = sum(abs(glx-sx),3);
gdifffy = sum(abs(gly-sy),3);
gdiffz = sum(abs(glz-sz),3);
```

která je následně sloučena do jedné hodnoty  $C_{GRAD}$  případně  $C_{SSIM}$

```
gdiff = gdifffx+gdifffy+gdiffz;
```

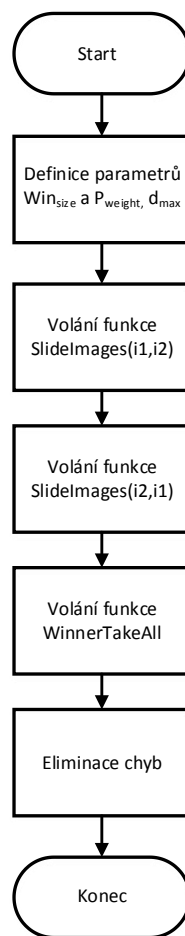
V následujícím kroku je provedena agregace míry podobnosti v okně dané velikosti pomocí funkce *imfilter*. V dalším kroku je proveden součet daných průměrovaných hodnot v barevném a gradientním prostoru s uvážením váhy gradientního obrazu a je získána celková hodnota míry podobnosti

```
d = CSAD+weight*CGRAD;
```

Získaná míra je pro každý pixel porovnána s hodnotou v matici *diff*, která uchovává do té doby nejnižší míru podobnosti (cenu disparity). Pokud je pro daný pixel a disparitu hodnota míry podobnosti menší než ta v matici, tak ji v matici nahradí a do matice disparit *disp* je uložena aktuální disparita jako ta nejvýhodnější. Pokud je pro daný pixel a disparitu hodnota větší než ta v matici *diff*, tak se hodnota v maticích *diff* ani *disp* nemění.

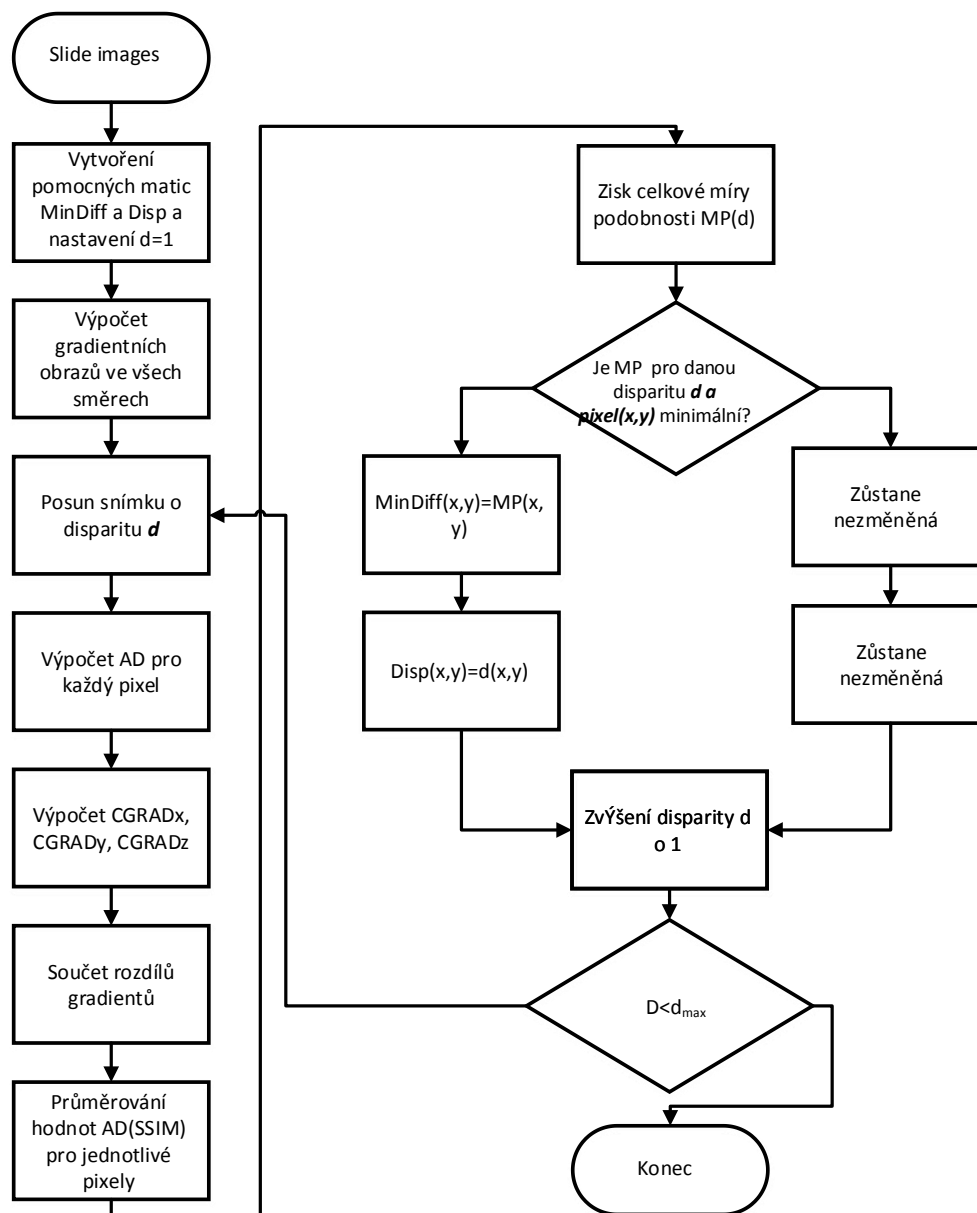
```
idx = find(d<mindiff); % test ceny aktuální ceny disparity
disparity(idx) = abs(i); %nahrazení doposud nejvhodnější disparity
mindiff(idx) = d(idx); %nahrazení ceny doposud nejvhodnější disparity
```

Popsaný postup je pomocí funkce *for* vykonán pro všechny zkoumané disparity a výstupem funkce je konečná podoba matic *diff* a *disp*.



Obr. 6.1: Vývojový diagram implementovaného postupu.





Obr. 6.2: Vývojový diagram funkce SlideImages zajišťující výpočet míry podobnosti a jejich agregaci.

## 6.2 Semiglobální metoda

Implementovaná lokální metoda lze použít jako základ pro semiglobální či globální metodu. Za tímto účelem, je v rámci lokální metody vytvořena trojrozměrná proměnná *DSI*. Rozměry proměnné jsou dány velikostí obrazu a parametrem  $d_{max}$  (maximální velikost disparity). Scanline je principiálně jednoduchý algoritmus vycházející z metrik lokální metody. Algoritmus využívá k určení disparity daného pixelu dvě informace: první informace je cena disparity získaná lokální metodou uložená v proměnné *DSI*, která informuje o podobnosti pixelů s pixelem v druhém obraze na pozici posunuté o danou disparitu. Druhá informace reprezentuje snahu o minimalizaci nespojitosti hloubkové mapy a je reprezentovaná druhým a třetím členem v následujícím vztahu

$$Dcont = normDSI(x, y, d) + abs(d - disp(x + 1, y)) + abs(d - disp(x - 1, y)), \quad (6.1)$$

kde  $x, y$  definují pozici v obraze,  $d$  je hodnota právě testované disparity,  $disp$  reprezentuje aktuální disparitní mapu a  $normDSI$  reprezentuje normovanou cenu disparity vypočtenou během lokální metody. Hodnota  $normDSI$  musí být normována na nižší hodnotu. Původní *DSI* totiž nabývá hodnot v řádu stovek a výše a musí být sníženo, aby nebyl zcela potlačen vliv druhé části rovnice, a který nabývá maximálně hodnoty rozdílu minimální a maximální disparity. Pro všechny body v obraze a všechny uvažované disparity je vypočtena hodnota  $E$  pomocí rovnice (6.1). Následně je nalezeno minimum  $E$  pro daný obrazový bod,

$$disp(x, y) = \min(E(x, y, :)). \quad (6.2)$$

Čímž je určena nová hodnota disparity, která je přiřazena do aktuální disparitní mapy, algoritmus lze provádět iteračně.

Algoritmus lze popsat vývojovým algoritmem na obrázku 6.3a. Byly provedeny experimenty s daným algoritmem:

- Výpočet ceny spojitosti ve větším než jednotkovém okolí bodu,
- Výpočet ceny nespojitosti nejen v řádku, ale i ve sloupci,
- Výpočet nespojitosti ne v hloubkové mapě, ale barevném obraze.

## 6.3 Belief propagation

Metoda Belief Propagation je popsána v teoretické části práci 5.3.2. V této kapitole se budeme věnovat popisu její konkrétní implementace v prostředí MATLAB. Implementace vychází z C++ implementace algoritmu [30] a mírně ji upravuje. Implementace může být popsána vývojovým algoritmem na obrázku Obr. 6.3b **Chyba! Nenalezen zdroj odkazů..** V prvním kroku je potřeba definovat následující vstupní parametry:

- `BP_ITERATIONS`: definuje počet iterací,
- `LABELS`: definuje počet labelů a určuje tak maximální uvažovanou disparitu
- `LAMBDA`: konstanta použitá ve funkci *SmoothnessCost* pro násobení rozdílu disparit.
- `SMOOTHNESS_TRUNC`: práh použitý ve funkci *SmoothnessCost*,

- `win_size_SAD`: velikost okna použitého v lokální metodě pro průměrování.
- `weight_SAD`: váha gradientní metriky v lokální metodě.

Nejdůležitější proměnnou implementace je trojrozměrná proměnná *Grid*. Proměnná *Grid* má tolik řádků, kolik má vstupní obraz pixelů. Následně má tato proměnná 5 sloupců, první čtyři slouží k uložení zprávy zasílané do 4 různých směrů a poslední uchovává cenu disparity pro daný obrazový bod a danou disparitu. Třetí rozměr proměnné je daný počtem uvažovaných disparit, tedy hodnotou proměnné *LABELS*.

V druhé fázi je vypočtena cena jednotlivých disparit a vstupní disparitní mapa, tento výpočet používá stejný postup jako lokální metoda popsaná výše. Výchozí disparitní mapa je tedy shodná s disparitní mapou získanou lokální metodou, implicitně je nastaveno použití metriky absolutních diferencí (AD).

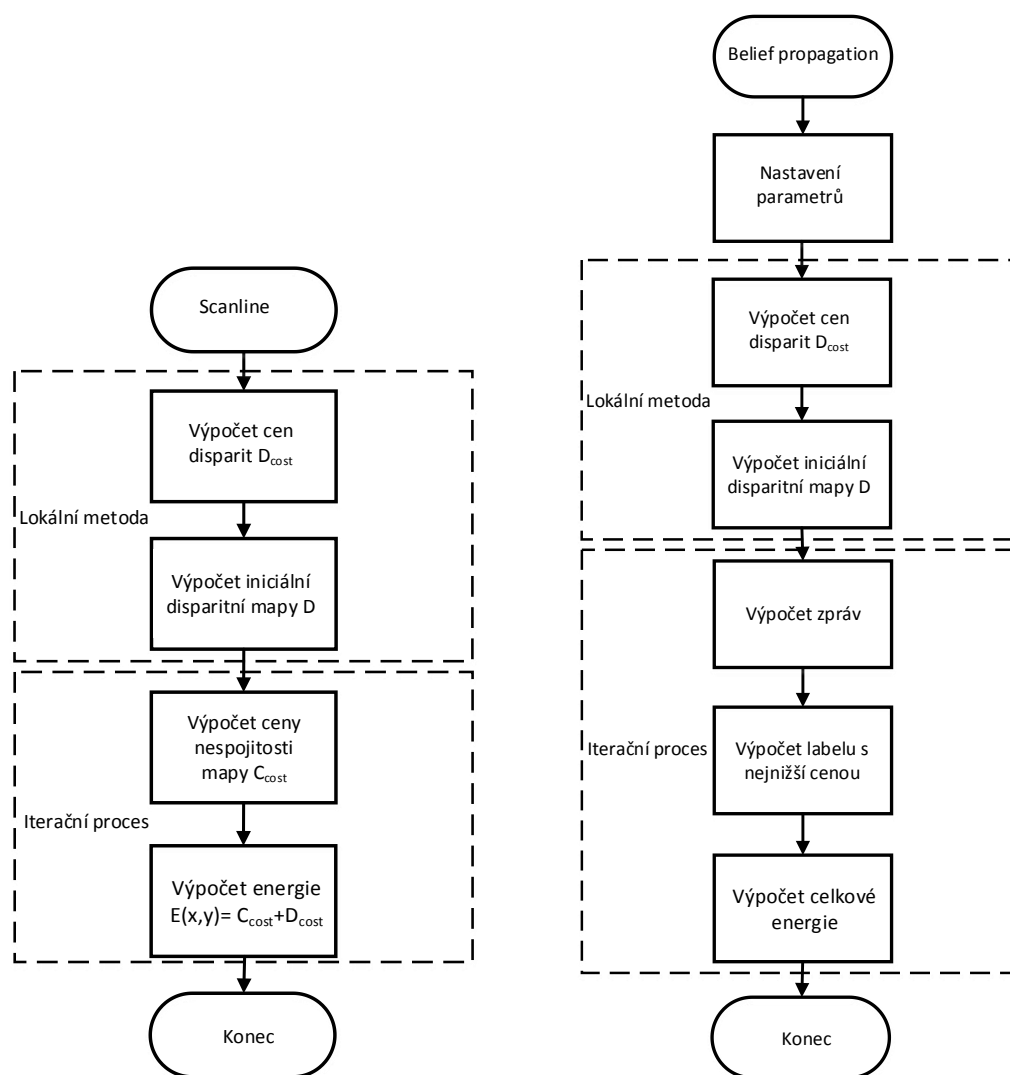
Následující fáze probíhá opakovaně (iterativně), a to do buď do okamžiku, kdy proběhne definovaný počet iterací (*BP\_ITERATION*) nebo do okamžiku, kdy přestane minimalizovaná proměnná *Energy* reprezentující energii dané hloubkové mapy klesat. V iterační části jsou provedeny základní kroky belief propagation popsané v kapitole 5.3.2. Pro šíření zpráv z každého uzlu (obrazový bod) do různých směrů je volána funkce *SendMsg*. Funkce počítá pro každý uzel (obrazový bod) minimální cenu-energii (Min-Sum BP) pro všechny možné disparity. Výsledná minimální cena je uložena pro danou disparitu  $d$ . Následně tato minimální cena je použita pro updatování zprávy pro sousední uzly (obrazové body).

Pro vytvoření zpráv posílaných z jednotlivých pixelů do čtyř různých směrů slouží volání čtyř velmi podobných funkcí: *SendMsgRIGHT*, *SendMsgLEFT*, *SendMsgUP*, *SendMsgDOWN*. Podrobněji proto bude popsána jen první z nich. Důležitou výhodou je skutečnost, že výpočet zprávy daným směrem je proveden pro všechny pixely obrazu současně (na rozdíl od původní C++ implementace). Tímto způsobem je efektivně využito možností prostředí MATLAB, a naopak je eliminována jeho velká nevýhoda v pomalém provádění cyklu *for*. Z teorie je zřejmé, že informace se skládá ze tří základních částí:

- součet zpráv příchozích z ostatních směrů,
- cena dané disparity,
- cena spojitosti.

Příchozí zprávy jsou uloženy v proměnné *Grid* a jsou sečteny 3 ze 4, neuvažuje se totiž zpráva, která přišla ze směru, do kterého bude vytvářena zpráva zaslána. Cena disparity je uložena v pátém sloupci proměnné *Grid*. Cena spojitosti se vypočítá pomocí volání funkce *SmoothnessCost*. Vstupy této funkce jsou hodnota uvažované disparity  $d$  a aktuální posunutá disparitní mapa *disp*. Funkci lze nejlépe popsat jednoduchým vývojovým diagramem na obrázku 8a. Funkce porovnává hodnotu disparity  $d$  a hodnotu disparitní mapy sousedního pixelu, a pokud pro daný obrazový bod rozdíl přesáhne hodnotu parametru *SMOOTHNESS\_TRUNC*, pak je výstup roven *SMOOTHNESS\_TRUNC*. Pokud nepřesáhne, tak je výstupní hodnota rovna rozdílu disparity  $d$  a hodnoty disparitní mapy pro daný obrazový bod. Funkce je prováděna současně pro všechny obrazové body. Součet těchto tří složek zprávy je proveden postupně pro všechny uvažované hodnoty disparity a jako odeslaná zpráva pro jednotlivé pixely je nakonec vybrána zpráva s nejmenší velikostí.

Obdobným způsobem jsou pro každý obrazový bod (pro všechny současně) určeny zprávy vysílané všemi čtyřmi směry. V posledním kroku iteračního procesu je pomocí volání funkce *MAP* přiřazena každému obrazovému bodu disparita, která minimalizují globální energii. Přiřazení je provedeno následujícím způsobem: Pro každý obrazový bod je určen součet příchozích zpráv a ceny disparity pro každou z disparit. Vzhledem k uložení těchto hodnot v proměnné *Grid* se vlastně pro každou disparitu jedná o součet všech sloupců této proměnné na dané hladině definované disparitou. Následně je určeno, pro kterou disparitu je součet, a tedy energie daného obrazového bodu, minimální a daná disparita je přiřazena danému obrazovému bodu. Následně začíná další iterace celého postupu a jsou znovu počítány vysílané zprávy.



Obr. 6.3: Vývojový diagram a) Semiglobální metody Scanline b) Globální metody Belief Propagation.

## 6.4 Metoda založená na nalezení výrazných bodů

Idea této metody je popsána v kapitole 5.5. Jednou větou lze metodu popsat následovně.

Hustá disparitní mapa může být vytvořena díky určitým předpokladům o jejích vlastnostech, ze znalosti horizontální disparity výrazných bodů scény. Celý postup testovaného přístupu lze popsat vývojovým diagramem na obrázku Obr. 6.4.

V prvním kroku algoritmu musí být tedy nalezeny korespondující body ve stereosnímkech. Pro nalezení výrazných bodů v obou obrazech a korespondenci mezi nimi je použita metoda SURF [29]. Samotné vyhledání výrazných bodů je provedeno pomocí funkce MATLABu

```
points1 = detectSURFFeatures(ilg, 'MetricThreshold', tresh);
```

, kde *ilg* je levý vstupní monochromatický snímek, 'MetricThreshold' je vlastnost definující hodnotu použité metriky SURF, pro kterou je bod označen za výrazný bod obrazu. *Thresh* je konkrétní hodnota výše popsaného prahu a uživatel ji definuje v uživatelském okně. Výstupní proměnná vrací informaci o nalezených bodech. Funkce je volána na levý i pravý snímek. Následně je volána funkce pro extrakci SURF deskriptorů pro nalezené body

```
[f1, vpts1] = extractFeatures(ilg, points1);
```

Deskriptory bodů z jednotlivých snímků jsou uloženy v proměnných *f1* a *f2*. V následujícím bodě jsou deskriptory bodů porovnány a jsou nalezeny korespondující páry.

```
indexPairs = matchFeatures(f1, f2);
```

Pomocí indexů jsou z původních skupin bodů *Points1* a *Points 2* vybrány pouze body z korespondujících párů a jsou získány proměnné *Pos1* a *Pos2*. Proměnné mají tolik řádků, kolik bylo nalezeno korespondujících párů a dva sloupce. V prvním sloupci je uložena vertikální pozice bodů a v druhém horizontální pozice.

Následuje rychlá eliminace falešných korespondencí na základě rozdílů vertikálních pozic bodů. Rozdíl vertikálních souřadnic v korespondujícím páru bodů by měl mít ideálně nulový vzhledem k tomu, že předpokládáme na vstupu stereosnímky. Z toho důvodu budou vyřazeny ty stereopáry, u nichž přesáhne vertikální disparita stanovený práh (experimentálně prozatím stanovený na 2).

V další fázi je určena pro nás zajímavá horizontální disparita korespondujících bodů, která je spočítána jako rozdíl horizontálních souřadnic bodů (tedy jako rozdíl druhých)

```
disparity=newPos1(:,1)-newPos2(:,1);
```

Pro přiřazení disparity jednotlivých bodů na příslušná místa disparitní mapy je vytvořena proměnná o stejné velikosti jako vstupní snímek

```
disp_image=zeros(x,y,1);
```

Do nově vytvořené proměnné jsou v cyklu *for* s iteračním proměnnou *j* vloženy disparity daných stereopárů na pozice odpovídající pozici v levém snímku.

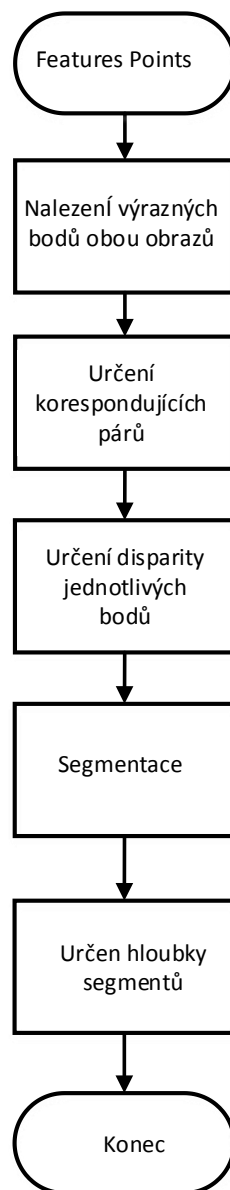
```
dsp_image(round(newPos1(j,2)),round(newPos1(j,1)))=disparity(j,1);
```

V druhé fázi je pro každý pár korespondujících bodů proveden následující postup. Pozice bodu v levém obraze je použita jako semínko jednoduché segmentace růstem ze semínka. Výstup metody je zobrazen na obrázku 9. Výstupem je oblast v okolí bodu s podobným jasnem. Následně jsou nalezeny indexy obrazových bodů snímku, které patří do sady nalezených korespondujících bodů a současně leží ve vysegmentované oblasti

```
idx_over=find(disp_image.*output~=0);
```

Tento krok je proveden z toho důvodu, že je možné, že oblast vzniklá kolem jednoho korespondujícího bodu bude obsahovat ještě další korespondující body, pro které také známe disparitu. Následuje cyklus for testující, jestli bod patřící k aktuálnímu segmentu již není přiřazen jinému segmentu. Jestliže není, tak je pixel označen pořadovým číslem segmentu. Jestli už patří k jinému segmentu, tak mu zůstane. Toto je jedno ze dvou míst algoritmu, kde by bylo možné testovat použití různých strategií.

Posledním krokem je výpočet disparity daného segmentu. Postupně je sečtena hloubka všech bodů segmentu, jehož disparitu známe. Následně jsou použity dva různé přístupy. V prvním z nich je hloubka všech bodů určena jako průměr známých disparit a v druhém případě jako jejich medián. V tomto místě postupu je také možné volit různé postupy.



Obr. 6.4 Vývojový diagram experimentální metody využívající výrazné body obrazu.

## 7 VÝSLEDKY

Kapitola pojednávající o dosažených výsledcích bude rozdělena do tří podkapitol. V první části bude testován vliv parametrů jednotlivých implementovaných metod (lokální metoda, Scanline, globální metoda Belief propagation, metoda s využitím výrazných bodů) na dosaženou kvalitu disparitní mapy. Informace o dosažené kvalitě disparitní mapy bude získána pomocí jejího porovnání s pravdivou disparitní mapou. Na základě prvního experimentu bylo vybráno ideální nastavení jednotlivých algoritmů pro konkrétní snímky a následně byly porovnány metody navzájem mezi sebou (7.2). Třetí podkapitola obsahuje ukázkou výsledných disparitních map. Testovací algoritmus používá jako základ zjednodušenou implementaci metody publikované v článku [20]. V testu byly použity obrázky ze stereobáze z university Middlebury [32]. Databáze obsahuje pravdivou disparitní mapu scén, což je zásadní fakt pro vyhodnocování kvality navržených algoritmů a vytvořených disparitních map. Tato databáze je velmi často používána a mnoho stávajících i nových algoritmů je s její pomocí evaluováno, což je velkou výhodou.

Kvalita disparitní mapy byla určována pomocí porovnání s pravdivou disparitní mapou. Výsledný parametr je v rozsahu 0-1, přičemž 0 znamená absolutní shodu a 1 naopak velmi nekvalitní hloubkovou mapu. Výpočet chyby disparitní mapy lze popsat následujícím pseudokódem:

```
1           $Error = |depth(x, y) - trueDepth(x, y)|$ 
2          If  $Error > \tau$  and  $trueDepth \neq 0$ 
3              {
4                   $n\_Error = n\_Error + 1$ 
5              }
6          If  $trueDepth \neq 0$ 
7              {
8                   $n\_Total = n\_Total + 1$ 
9              }
10          $d\_error = \frac{n\_Error}{n\_Total}$ 
```

Je vypočítán rozdíl mezi disparitou konkrétního pixelu na dané pozici obrazu definované jeho horizontální (x) a vertikální souřadnicí (y). Jestliže absolutní hodnota tohoto rozdílu překročí stanovený práh a současně má pravdivá disparitní mapa na dané pozici nenulovou hodnotu, pak je zvýšen počet chybových pixelů  $n\_Error$  o jedna. Současně je spočítán celkový počet pixelů  $n\_Total$  s definovanou (nenulovou) disparitou. Poměr těchto dvou hodnot stanovuje kvalitu (celkovou chybovost) hloubkové mapy.

### 7.1 Parametry lokální metody

Ve vytvořené aplikaci je použito několik takových funkcí popsaných již v kapitole 5.1: SAD a SSIM a výhodnost jejich použití byla testována. Dalším objektem experimentu bylo testování vlivu velikosti vyhledávacího okna na přesnost hloubkové mapy a vliv váhy gradientního parametru. Krátké shrnutí výsledků je v tabulkách č. 7.1

a č. 7.2 a v grafech na obr. 7.3 a 7.4. Experiment byl proveden s využitím 4 snímků z databáze z university Middlebury [32]. Použité stereo snímky a jejich pravdivá hloubková mapy jsou zobrazeny na obrázku 7.1. Na obrázku 7.2 je zobrazeno několik vytvořených hloubkových map

Z dosažených výsledků je zřejmé, že kvalita disparitní mapy je relativně odlišná pro různé scény, nejhorší kvalita je jednoznačně dosažena pro snímek Mid, který obsahuje velkou konstantní plochu pozadí, se kterou si lokální metoda není schopná poradit.

Tabulka č. 7.1 obsahuje výsledky první části experimentu, kdy byl konstantní parametr váhy rozdílu gradientů  $P_{weight} = 9$  a byla měněna velikost okna pro agregaci míry podobnosti pixelu  $win_{size}$ . Z výsledků je patrné, že kvalita hloubkové mapy se s velikostí okna mění. Pro některé scény a míru podobnost je závislost větší a kvalita hloubkové mapy se změní až o 5 procent (SSIM a snímek Wood), a naopak některým snímkům se kvalita disparitní mapy vlivem použití okna různých velikostí takřka nezmění. Nicméně není možné stanovit jednu správnou velikost okna napříč všemi testovanými obrazy. Je zřejmé, že kvalita hloubkové mapy a optimální velikost použitého okna závisí do velké míry na obsahu scény. Stejně tak není ani možné z testovaného vzorku říct, zda je lepší použít míru podobnosti AD nebo SSIM.

Tabulka 7.2 obsahuje výsledky druhého experimentu, kdy je nastavena velikost okna na 12 pixelů a mění se velikost parametru  $P_{weight}$  v rozsahu od 3 do 27. Z výsledků je tentokrát zřejmé, že nejlepší kvality disparitní mapy je dosahováno při nízkých hodnotách tohoto parametru. Nepřímá úměrnost platí pro obě použité metriky a pro všechny snímky.

Z dosažených výsledků je zřejmé, že pokud je použita lokální metoda, tak je důležité, aby finální aplikace, jenž vznikne v rámci bakalářské práce, umožňovala měnit parametry metody.

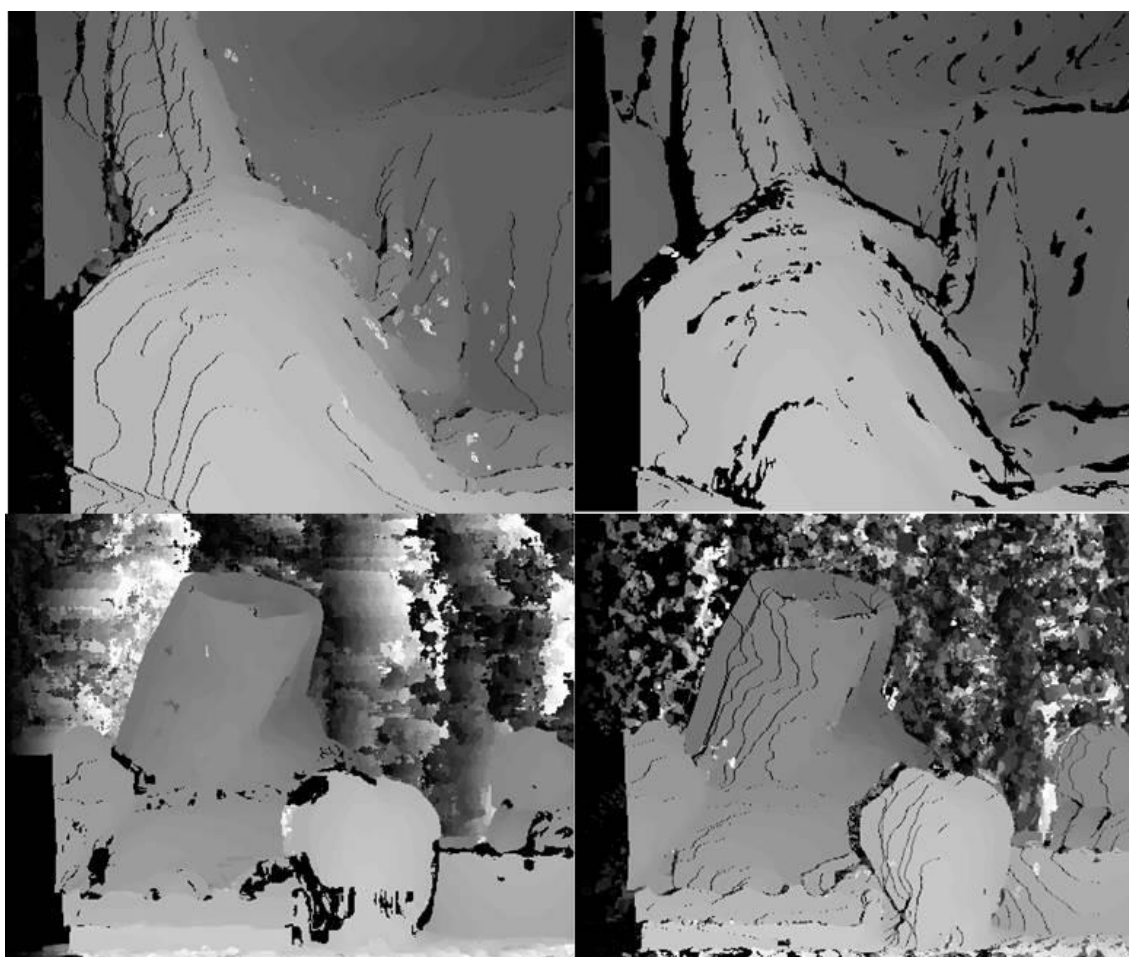
Tab. 7.1: Výsledky experimentu testujícího vliv použité míry podobnosti a vliv velikosti okna pro agregaci míry podobnosti.

	Velikost okna	3	6	9	12	15
Wood	AD	0,1918	0,1848	0,1784	0,1823	0,1861
	SSIM	0,2207	0,1629	0,1561	0,1610	0,1636
Baby	AD	0,2251	0,2079	0,1984	0,1927	0,1900
	SSIM	0,2554	0,2063	0,1935	0,1862	0,1863
Cloth	AD	0,1614	0,1548	0,1497	0,1472	0,1454
	SSIM	0,1756	0,1801	0,1916	0,2069	0,2233
Mid	AD	0,5022	0,4861	0,4726	0,4641	0,4588
	SSIM	0,5022	0,4807	0,4737	0,4789	0,4860

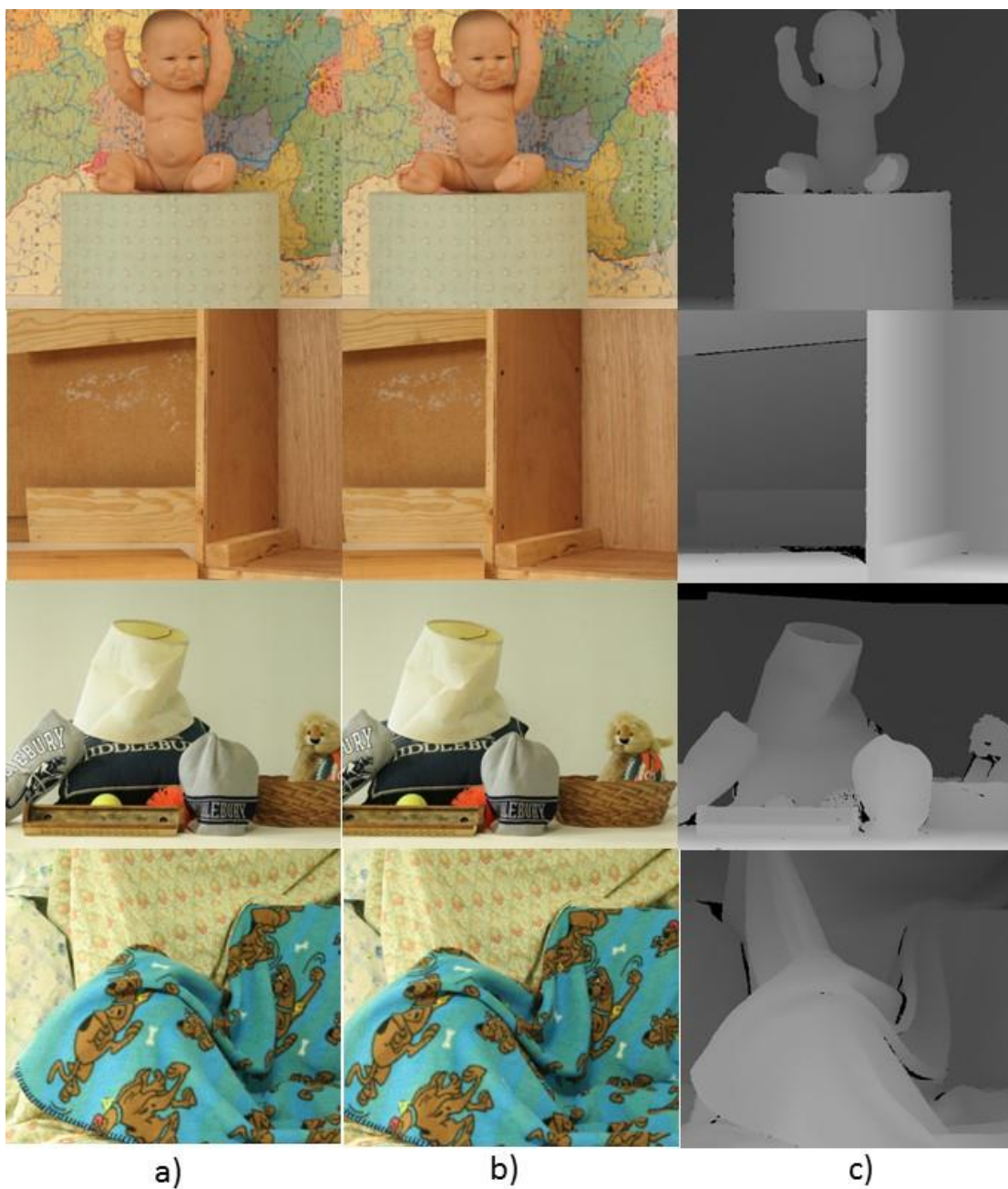


Tab. 7.2: Výsledky experimentu testujícího vliv použité míry podobnosti a vliv váhy míry podobnosti spočítané v gradientním prostoru.

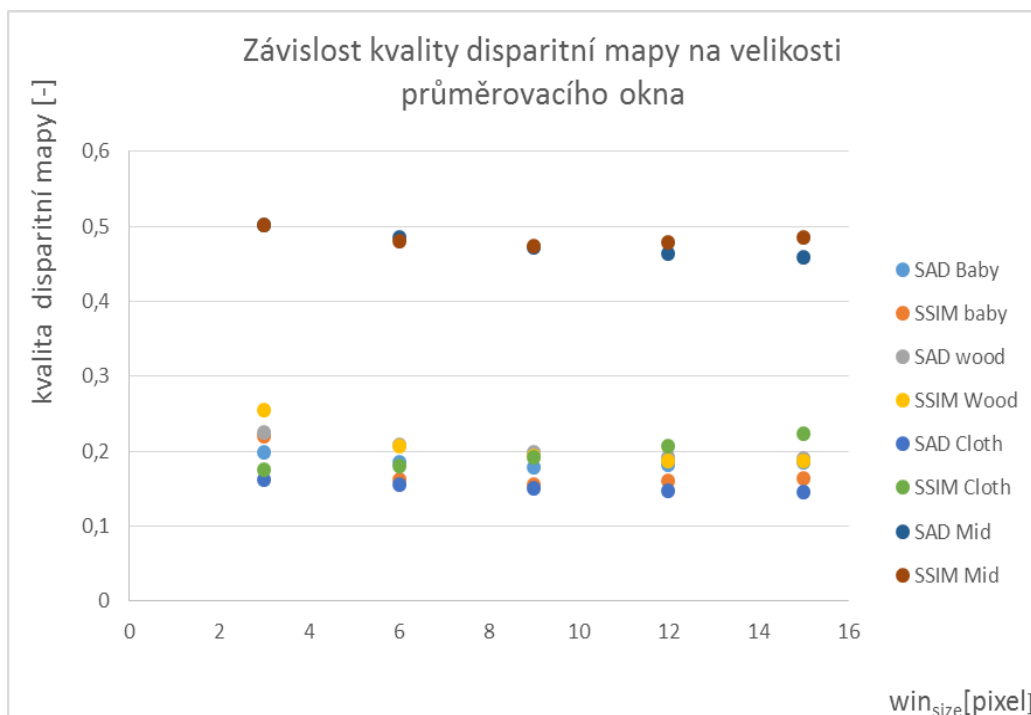
	Velikost váhy	3	9	15	21	27
Wood	AD	0,1765	0,1773	0,1784	0,1788	0,1792
	SSIM	0,1537	0,1537	0,1561	0,1614	0,1703
Baby	AD	0,1940	0,1975	0,1984	0,1988	0,1992
	SSIM	0,1445	0,1559	0,1935	0,2473	0,2989
Cloth	AD	0,1480	0,1491	0,1497	0,1498	0,1499
	SSIM	0,1187	0,1384	0,1916	0,2657	0,3291
Mid	AD	0,4688	0,4712	0,4725	0,4732	0,4737
	SSIM	0,4425	0,4576	0,4737	0,4889	0,5066



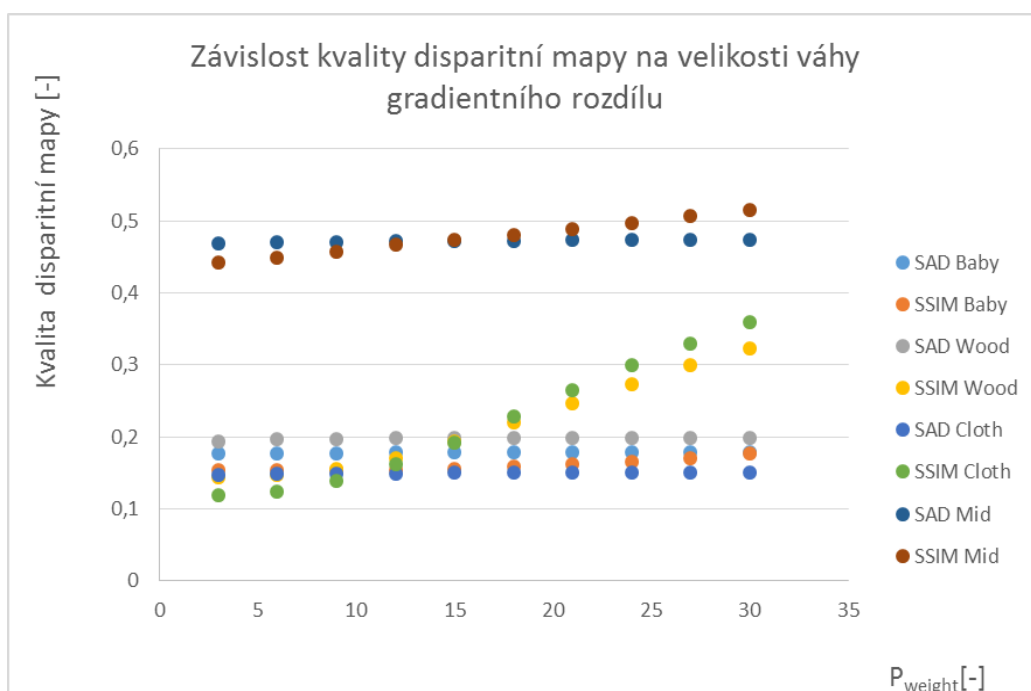
Obr. 7.1: Výstupní hloubkové mapy získané během experimentu V první řádce snímek Cloth a ve druhém Mid. V prvním sloupci použití metriky SSIM a v druhém SAD.



Obr. 7.2: Obrázky z databáze [32] použité v provedeném experimentu a) Levý snímek b) Pravý snímek c) pravdivá hloubková mapa. První řada shora scéna Baby, druhá řada scéna Wood, třetí řada scéna Mid a čtvrtá řada scéna Cloth.



Obr. 7.3: Závislost kvality disparitních map získaných různými mírami podobnosti na velikosti okna použitého pro agregaci.



Obr. 7.4: Závislost kvality disparitních map získaných různými mírami podobnosti na váze míry podobnosti spočítané v gradientním prostoru.

## 7.2 Porovnání implementovaných metod

V následující kapitole budou vzájemně porovnány výsledky dosažené použitím různých metod. Porovnávaná bude dosažená kvalita vypočítaných disparitních map. Na základě experimentu, jehož průběh a výsledky jsou popsány v kapitole 7.1, byly stanoveny parametry lokálních metod. Pro globální metody a pro metodu založenou na hledání výrazných bodů bylo provedeno jednoduché testování vlivu použitých parametrů a pro porovnání metod mezi sebou byly zvoleny parametry, pro které vycházely kvalitní výsledky, přestože o experimentu nepojednává samostatná podkapitola. Výsledky dosažené pěti různými vlastními implementacemi metod:

- lokální metoda s metrikou AD (dále značeno jen jako SAD),
- lokální metoda s metrikou SSIM (SSIM),
- semiglobální metoda s optimalizací Scanline (SL),
- globální metoda s optimalizací metodou Belief Propagation (BP),
- metoda založená na hledání výrazných bodech obrazu (FPoints),

byly současně porovnány s dalšími dvěma disparitními mapami. První z nich byla získána staženou implementací metody využívající optimalizační algoritmus Adaptive Random Walk [33], druhou z nich je metoda, která je součástí prostředí MATLAB [34] (nadále označení MATLAB). Využité metody budou stručně popsány.

Funkce provádějící výpočet metodou RandomWalk (dále značeno jen jako RW) má následující hlavičku

```
function [ left_disparity_map, right_disparity_map ] =  
stereo_matching( iL, iR, maxD, sigma_e, tau_e, sigma_psi, tau_psi,  
sigma_g, tau_g, sigma_c, tau_c, r, t, superpixel_size, spatial_weight,  
lookup_table, penalty_function )
```

Z počtu vstupů je patrné, že se jedná o komplexní, relativně složitý pokročilý přístup. Vzhledem k tomu, že se jedná o přejatý kód použitý pouze pro srovnání výsledků, nebudeme se zabývat jeho postupem ani vysvětlením významu jednotlivých vstupů. Hodnoty vstupů byly ponechány na defaultně autorem nastavených hodnotách a vliv jejich změny nebyl zkoumán. Experiment s vlivem parametrů by byl jistě přínosný, ale přesahuje rozsah bakalářské práce a zamýšleného použití metody pouze pro zběžné srovnání s vlastními dosaženými výsledky.

Funkce pro výpočet disparitní mapy, jenž je součástí prostředí MATLAB, je založena na lokální metrice AD a má semiglobální nastavbu. Podle předpokladů by tedy měla dosahovat podobných výsledků, jako má lokální implementace, popřípadě jako rozšíření s algoritmem Scanline. Funkce může být volána následujícím způsobem

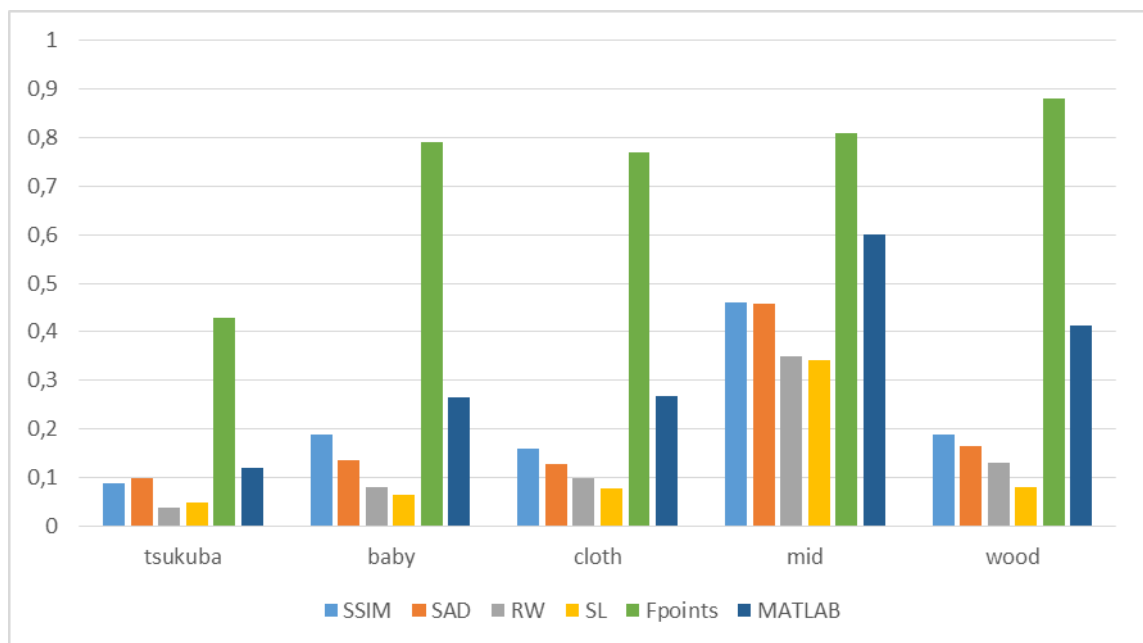
```
disparityMap = disparity(i1,i2),
```

v nejjednodušším případě jsou tedy vstupy funkce pouze levý a pravý snímek a parametry metody jsou ponechány na defaultních hodnotách. V našem případě jsme použili právě defaultní nastavení s výjimkou nastavení rozsahu disparit pomocí parametru *DisparityRange*. Dalšími parametry jsou velikost bloku pro optimalizaci a několik prahů používaných v rozhodování.

Všechny metody byly testovány na několika snímcích. Výsledky jsou přehledně shrnuty v tabulce 7.3 a vyneseny do grafu na obrázku 7.5.

Tab. 7.3: Porovnání implementovaných metod, přejaté metody a funkce prostředí MATLAB.

	SAD	SSIM	SL	BP	FPoints	RW	MATLAB
<b>Wood</b>	0,1636	0,1867	0,0796	0,1781	0,8809	0,1298	0,4127
<b>Baby</b>	0,1349	0,1872	0,0633	0,1250	0,7898	0,0803	0,2656
<b>Cloth</b>	0,1269	0,1595	0,0773	0,1562	0,7710	0,0968	0,2686
<b>Mid</b>	0,4590	0,4614	0,3419	0,4580	0,80086	0,3500	0,6002
<b>Tsukuba</b>	0,0987	0,0869	0,0489	0,0734	0,4291	0,0374	0,1190



Obr. 7.5 Závislost kvality vypočítané disparitní mapy metodě pro různé snímky.

Z výsledků lze odvodit několik skutečností. Pouze potvrzením výsledků popsanych v kapitole 7.1 je fakt, že disparitní mapy jsou přibližně stejně kvalitní při použití různých metrik v lokální metodě. Nejlepších výsledků je v drtivé většině případů dosaženo při použití metody Scanline. Pouze občas dosáhne lepších výsledků přejatá implementace metody s optimalizací Random Walk. Metoda postavená na Belief Propagation je schopná vylepšit vstupní disparitní mapu, nicméně výstupní disparitní mapa nedosahuje kvality dvou výše zmíněných. Tato skutečnosti zcela neodpovídá předpokladů o kvalitních výsledcích dosahovaných touto metodou. Zdaleka nejhorších výsledků je dosaženo, pokud je využita metoda založená na výrazných bodech, což naopak není překvapivé. Metoda založená na výrazných bodech by pro lepší využití musela využívat zejména lepší segmentaci. Problémem disparitních map získaných touto metodou není ani tak přesnost určených disparit, ale to, že relativně u vysokého

počtu bodů není určena disparita vůbec. Tento problém je při dalším práci relativně snadno řešitelný dvěma různými způsoby. První z nich je využití lepší segmentace, protože je zřejmé, že problém je v tom, že segmenty rostoucí ze semínek výrazných bodů nepokryjí celý snímek. Druhým možným řešením je přiřazení disparit v dalším kroku pomocí nějakých pravidel, vycházejících například z předpokladu spojitosti hloubkové mapy. Výsledky dosažené metodou MATLABu jsou druhé nejhorší a jsou výrazně horší než výsledky dosažené prostou lokální metodou. Tento fakt znamená, že vytvořená implementace lokální metody je zdařilá. Lepších výsledků by snad bylo možné dosáhnout variováním parametrů metody.

### 7.3 Ukázka výsledků

V této kapitole bude zobrazeno několik vypočítaných disparitních map při použití různých metod a jejich parametrů. Vstupní levý a pravý snímek scény použité pro ukázkou dosažených výsledků jsou zobrazeny na obrázku Obr. 7.6. Pravdivá disparitní mapa dané scény je zobrazena na obrázku Obr. 7.7. Disparita scény Tsukuba nabývá hodnot v rozsahu 1-16. Disparitní mapy snímku Tsukuba získané lokálními metodami při nastavení různých parametrů jsou zobrazeny na obrázku Obr. 7.8.



Obr. 7.6: Vstupní levý a pravý snímek scény Tsukuba.

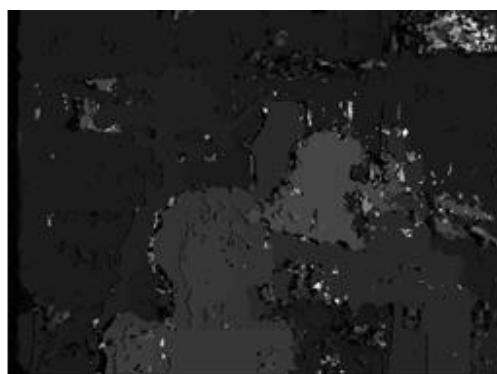


Obr. 7.7: Pravdivá disparitní mapa scény Tsukuba.

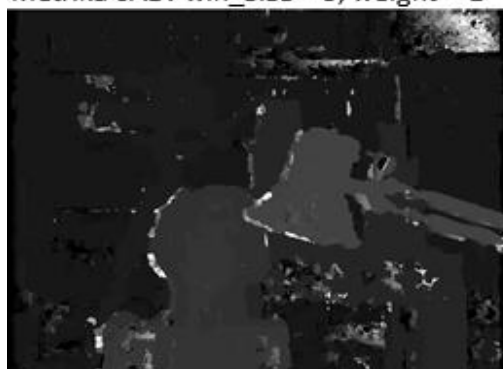




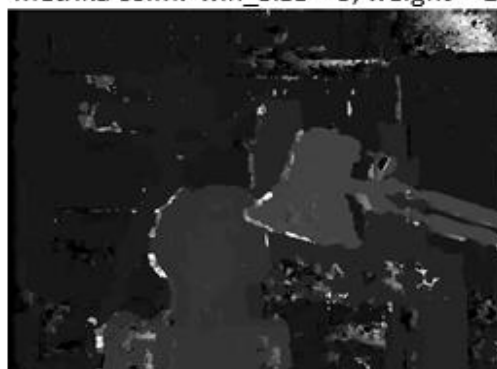
Metrika SAD: win\_size = 3, weight = 1



Metrika SSIM: win\_size = 3, weight = 1



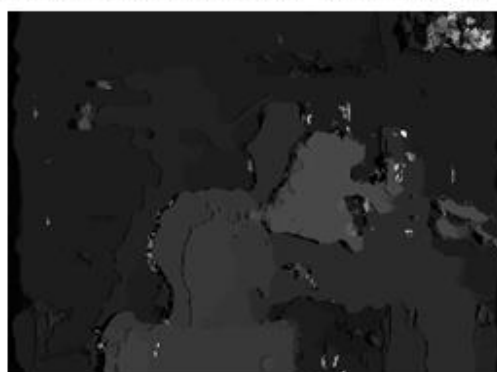
Metrika SAD: win\_size = 6, weight = 1



Metrika SSIM: win\_size = 6, weight = 1



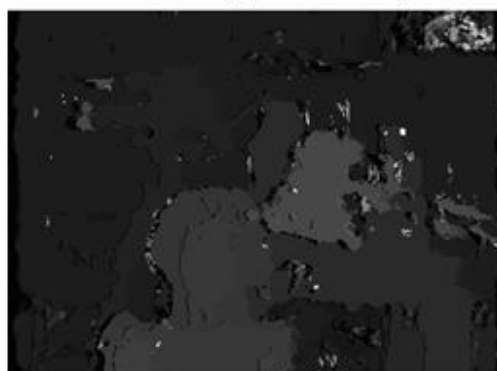
Metrika SAD: win\_size = 9, weight = 1



Metrika SSIM: win\_size = 9, weight = 1



Metrika SAD: win\_size = 15, weight = 1



Metrika SSIM: win\_size = 15, weight = 1

Obr. 7.8 Disparitní mapy snímku Tsukuba získané lokálními metodami při nastavení různých parametrů.

## 8 VÝSTUPY BAKALÁŘSKÉ PRÁCE

### 8.1 Databáze stereosnímků

V rámci bakalářské práce byla vytvořena databáze stereosnímků. Cílem bylo vytvořit databázi vlastních stereosnímků, která bude moci být v budoucnu využita pro testování algoritmů pro výpočet disparitních map. Z tohoto důvodu neobsahuje databáze pouze samotné stereo snímky, ale další informace. Databáze obsahuje:

- levý a pravý stereo snímek,
- hloubkovou mapu pro daný stereo snímek (u vytvořených zatiší),
- vlastnosti stereopáru (prostorovou aktivitu snímku, SSIM mezi snímky),
- vnitřní parametry kamery použité pro snímání (stereobáze  $B$ , ohnisková vzdálenost  $f$ ).

Snímky v databázi lze dělit podle obsahu scény. Jedná o dvě skupiny: naaranžované vnitřní scény s pravdivou hloubkovou mapou (20 snímků) a venkovní fotografie reálných scén (40 snímků). Snímky byly pořízeny fotoaparátem FUJIFILM FinePix Real 3D W3 se dvěma objektivy se stereobází  $B=6\text{cm}$  a ohniskovou vzdáleností  $f=35\text{mm}$ . Rozlišení každého ze snímků je  $1792 \times 2016$ . Pravdivá hloubková mapa byla získána pomocí zařízení Kinect. Hloubková mapa je uložena ve formě MATLAB souboru s příponou *mat*. Rozlišení hloubkové mapy je  $480 \times 640$ . Matice hloubkové mapy obsahuje v milimetrech vzdálenost prostorových bodů reprezentovaných danými obrazovými body od snímáče. Databáze je elektronickou přílohou bakalářské práce. Součástí přílohy bakalářské práce jsou také vypočítané parametry snímků. Databáze je doplněna o parametry snímků kvůli možnosti v budoucnosti zkoumat jaký vliv má obsah scény na úspěšnosti tvorby disparitní mapy. Parametr SSIM je počítán pomocí jednoduchého volně šiřitelného kódu. Základ výpočtu prostorové aktivity snímku je výpočet rozdílů mezi sousedícími obrazovými body. M-file výpočtu parametru je součástí elektronické přílohy práce.

### 8.2 Uživatelské rozhraní

Důležitou součástí bakalářské práce je vytvoření uživatelsky přívětivého grafického rozhraní. Po spuštění programu uživatel uvidí úvodní obrazovku programu, která je zobrazena na obrázku Obr. 8.2. Obr. 8.1 již včetně vypočítaných disparitních map. Na této obrazovce jsou uživateli dostupné všechny funkce programu a v průběhu používání na ní názorně uvidí téměř všechny dosažené výsledky. Uživatelské rozhraní umožňuje využití všech metod pro tvorbu disparitní mapy, jejichž implementace je popsána v kapitole 6. Jsou jimi tyto metody: lokální metoda založená na podobnosti okolí bodů, semiglobální metoda Scanline, globální metoda využívající optimalizaci Belief Propagation, metoda založená na hledání korespondujících bodů.

Dominantní část rozhraní slouží pro zobrazení vstupních levých a pravých snímků, pravdivé disparitní mapy a vypočítaných disparitních map. V levé části okna



jsou tlačítka sloužící pro otevření levého a pravého snímku zvlášť, což může být v určitých situacích velmi užitečné. Dalším tlačítkem *Open True Disparity Map* je možné otevřít a zobrazit pravdivou disparitní mapu daného snímku. Načtená pravdivá mapa může následně sloužit pro výpočet kvality vypočítané disparitní mapy. Z tohoto důvodu je tlačítko *Disparity Map Quality* zpřístupněno až po načtení pravdivé disparitní mapy. Důležitou součástí okna je volba mezi použitými metodami v panelu *Method Choice*, výběr použité metody ovlivňuje vzhled panelu v pravé části obrazovky a tedy parametry, které uživatel může nastavovat. Detailní pohled na panely pro nastavení parametrů jednotlivých metod je na obrázku Obr. 8.1. Po stisku tlačítka *Calculation* je volána funkce provádějící výpočet disparitní mapy zvolenou metodou. Za předpokladu, že má uživatel k dispozici pravdivou disparitní mapu příslušící k danému stereopáru, může po výpočtu hloubkové mapy provést porovnání získané disparitní mapy a pravdivé disparitní mapy. Uživatel zvolí, jakou disparitní mapu chce porovnávat pomocí skupiny *RadioButtonu* v panelu *Depth map Quality* a následně výpočet spustí stiskem tlačítka *Depth map Quality*.

Vytvořená aplikace také umožňuje vytvoření hloubkové mapy a prostorového modelu. Pro tuto funkci uživatel musí zvolit možnost *Depth map and 3D model* v panelu *Method Choice*. V nově zobrazeném panelu *Depth map and 3D model calculation* uživatel musí nastavit ohniskovou vzdálenost fotoaparátu  $f$  a stereo bázi  $B$  a následně výpočet spustí stiskem tlačítka *Calculation* v daném panelu.

Aplikace dále umožňuje uložení získané disparitní mapy ve formátu *png* a jako MATLAB soubor *mat*. Uložení je provedeno po stisku tlačítka *Save Map*, volba ukládané mapy je provedena pole *RadioButtonu* v poli *Depthmap quality*.

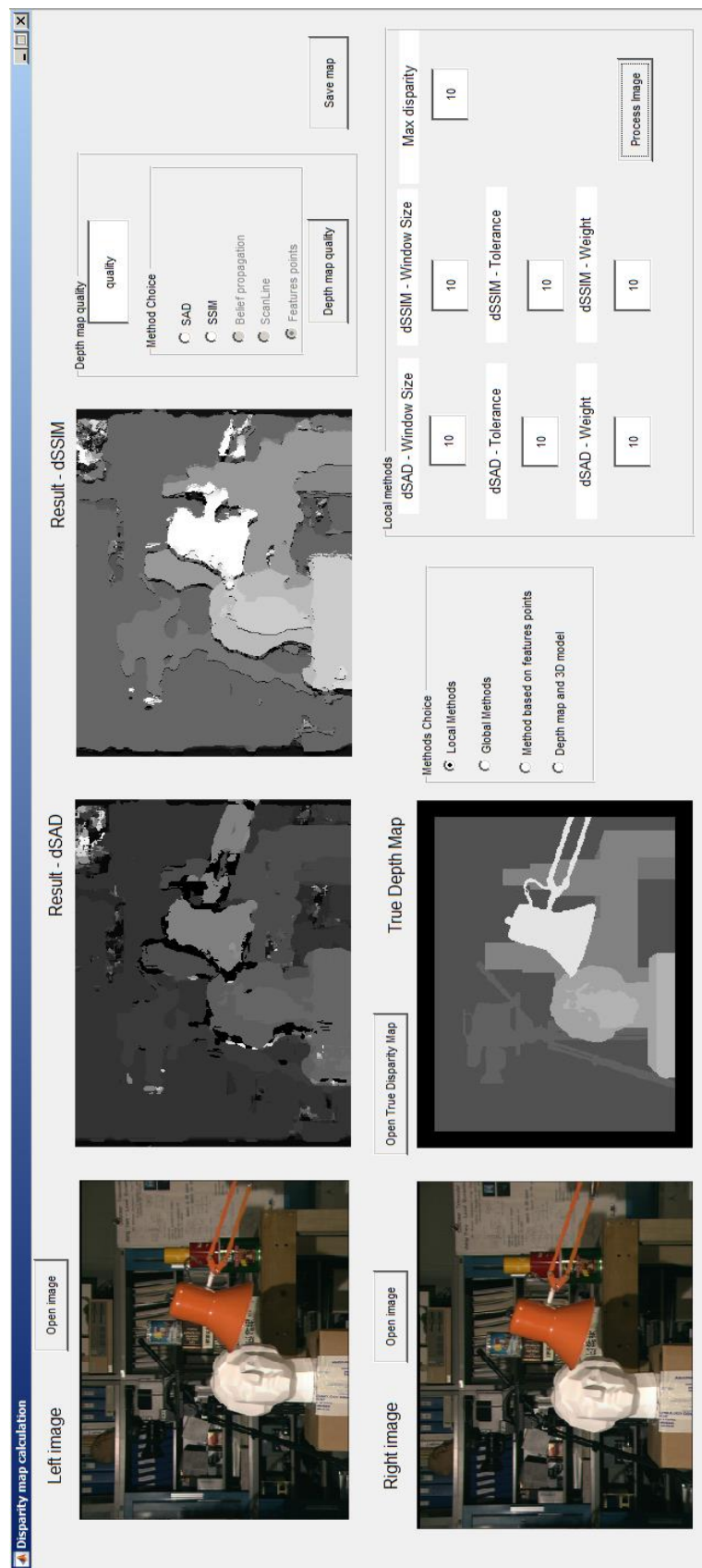
The image shows two panels for parameter settings. The top panel is titled 'Local methods' and contains the following controls:

- dSAD - Window Size:** Input field with value 10.
- dSSIM - Window Size:** Input field with value 10.
- Max disparity:** Input field with value 10.
- dSAD - Tolerance:** Input field with value 10.
- dSSIM - Tolerance:** Input field with value 10.
- dSAD - Weight:** Input field with value 10.
- dSSIM - Weight:** Input field with value 10.
- Process Image:** Button.

The bottom panel is titled 'Global methods' and contains the following controls:

- dSAD - Window Size:** Input field with value 10.
- Number of iteration:** Input field with value 10.
- Max disparity:** Input field with value 10.
- dSAD - Tolerance:** Input field with value 10.
- Lambda:** Input field with value 10.
- dSAD - Weight:** Input field with value 10.
- SMOOTHNESS\_TRUNC:** Input field with value 10.
- Process Image:** Button.

Obr. 8.1 Panely pro nastavení a) Lokálních metod b) Globálních metod.



Obr. 8.2 Uživatelské okno aplikace.

## 9 ZÁVĚR

Cílem bakalářské práce bylo seznámit se s různými metodami a přístupy tvorby disparitní mapy, vytvořit aplikaci pro tvorbu disparitní mapy a vytvořit jednoduchou databázi stereosnímků. Všechny cíle bakalářské práce se podařilo splnit. Disparitní mapa je šedotónový obraz reprezentující prostorové uspořádání snímané scény. Disparitní mapa je důležitá zejména pro 3D TV, a to z několika důvodů. Disparitní mapa může být použita pro vytvoření nových virtuálních pohledů na scénu, z tohoto důvodu je součástí několika formátů pro přenos 3D videa. Disparitní mapa může být získána několika způsoby. První z nich je aktivní způsob, který spočívá ve využití přídavné informace přidané do scény zpravidla ve formě projekce určité informace. Zástupcem aktivních metod může být například zařízení KINECT s využitím infra zářiče a profilometrie. Druhý způsob je založený na snímání scény více kamerami, typickým zástupcem je stereometrie se snímáním scény dvěma kamerami, jejichž horizontální pozice se liší o sterobázi. Bakalářská práce se zabývá výlučně pasivní stereo metodou zisku hloubkové mapy.

V první části bakalářské práce je popsán samotný mechanismus stereopsie a stereoskopie, z něhož vyplývá důležitost disparitní a hloubkové mapy, na jejichž základě pracují v dnešní době všechny 3DTV systémy. Následně jsou popsány všechny používané a dostupné systémy 3D televize a 3D projekce. V následující kapitole jsou definovány nejdůležitější pojmy, které se týkají dané oblasti, a jenž se v práci často používají. Tvorba disparitní mapy patří mezi metody zpracování obrazu a je to v rámci zpracování obrazu pokročilá metoda spadající do závěrečné fáze zpracování obrazu, tzv. pochopení obsahu obrazu. Z tohoto důvodu jsou během tvorby disparitní mapy použity různé metody zpracování obrazu. Proto jsou tyto metody stručně popsány v kapitole 3. V této metodě je z důvodu celistvosti práce popsán i způsob samotné reprezentace scény obrazem, tedy barevné modely či šedotónový obraz.

V další kapitole je popsán princip hledání bodových korespondencí, který je fundamentálním krokem tvorby disparitních map. Dále jsou zde metody pro tvorbu disparitní mapy rozděleny na metody lokální a globální. Jednotlivé kroky těchto skupin metod jsou popsány a vysvětleny včetně několika základních funkcí pro výpočet míry podobnosti obrazových bodů. V kapitole jsou následně podrobněji popsány některé často používané metody globálních algoritmů, jako je Belief Propagation, Graph Cut nebo Skyline. Výsledkem studia různých metod bylo stanovení postupu v další etapě bakalářské práce a volba algoritmů, které byly implementovány v aplikaci napsané v prostředí MATLAB.

Následující kapitola se věnuje samotné tvorbě aplikace pro tvorbu hloubkových map. V kapitole jsou popsány čtyři použité metody, pro názorný popis je vždy použit vývojový diagram. Jako první byl implementován základní algoritmus lokální metody pro tvorbu disparitní mapy. Jako míra podobnosti byly použity metriky SAD a SSIM pospané v kapitole 5.1. Následují semiglobální metoda Scanline a globální metoda využívající algoritmus Belief Propagation. Základem obou metod je disparitní mapa získaná lokální metodou, která je následně optimalizována z hlediska minimalizace energie. Zjednodušeně řečeno se jedná o minimalizaci ceny disparit a ceny spojitosti disparitní mapy. Poslední metodou je metoda založená na nalezení výrazných bodů obrazů. Tato metoda je z větších částí experimentální a bylo by potřeba dále testovat

použití různých postupů. Zejména použití různých způsobů segmentace obrazu.

Výsledky dosažené ve vytvořené aplikaci použitím různých metod jsou popsány v další kapitole s názvem Výsledky. Testování vytvořené aplikace bylo provedeno na pěti snímcích. Byly vybrány snímky s relativně odlišným obsahem scény. Z výsledků je patrné, že kvalita hloubkové mapy je do značné míry právě na obsahu scény závislá. Z toho důvodu by nebylo možné zcela vyjádřit funkčnost implementací ani při mnohem větším počtu snímků. Nicméně pro základní porovnání metod a zkoumání vlivu parametrů a zejména pro prokázání funkčnosti implementace je daný počet snímků dostačující. Kapitola obsahuje zejména porovnání výsledků lokálních metod při nastavení různých parametrů. Byl otestován vliv použití různé metriky a různé velikosti okna na kvalitu výsledné disparitní mapy. V experimentu byly použity snímky z databáze, u kterých byla dostupná pravdivá disparitní mapa, se kterou mohl být porovnán výsledek a stanovena tak kvalita vytvořených disparitních map. Výsledky jsou shrnuty v tabulce a ve formě grafů. Z výsledků je zřejmé, že nejde jednoznačně říci, jestli je lepších výsledků dosahováno při použití metriky absolutních diferencí nebo SSIM. Z výsledků lze vyznívat, že disparitní mapy lepší kvality jsou získány zpravidla při použití většího průměrovacího okna, a to zejména u metriky absolutních diferencí. Tato skutečnost odpovídá teorii. Naopak je dosahována lepší kvality výsledných map při použití spíše menších vah diferenciálních měr. Na druhou stranu jen při absolutním vyloučení použití diferenciálních obrazů je vždy výsledná kvalita disparitní mapy horší. Velký vliv na kvalitu disparitní mapy mají parametry daného snímku. V další části experimentu byly výsledky získané jednotlivými metodami. Z výsledků je patrné, že použití optimalizačních metod Scanline a Belief Propagation vždy zlepši kvalitu disparitní mapy. Větší zlepšení vykazuje algoritmus Scanline, což vzbuzuje myšlenku, že v implementaci algoritmu Belief Propagation je nějaká menší chyba, která brání v optimálním běhu algoritmu, ta se však nepodařila odhalit.

Důležitou součástí vytvořené aplikace je uživatelské rozhraní, o kterém pojednává samostatná kapitola, která je z části příručkou jak aplikaci používat a z části popisem jak aplikace jako celek funguje. Vše, co uživatel potřebuje, se nachází v jednom jediném okně aplikace, což je uživatelsky příjemné. Snahou bylo udělat aplikaci pro uživatele přehlednou. Uživatelské rozhraní nabízí navíc kromě samotné tvorby disparitních map také přepočítání disparitní mapy na hloubkovou mapu a na prostorový model scény. Poslední funkcí je možnost porovnat vytvořenou disparitní mapu s pravdivou disparitní mapou snímku, pokud je pravdivá mapa k dispozici. Výsledné disparitní mapy mohou být uloženy ve formátu *png*.

Posledním bodem zadání bakalářské práce bylo vytvoření databáze rektifikovaných stereosnímků. K vytvoření stereosnímků byl použit stereofotoaparát, proto nebyla nutná rektifikace snímků. Databáze obsahuje dva typy scén. Každý snímek scény obsahuje informaci o vnitřních parametrech použité kamery a parametry snímku. Některé snímky obsahují také pravdivou hloubkovou mapu scény pořízenou zařízením Kinect.

## LITERATURA

- [1] SCHINDLER R.. *Zrakové vady – XVII. Binokulární vidění* [online]. Praha: Tyfloservis SONS, ©1994–2012 [cit. 1. listopadu 2015]. Dostupné na www: < <http://www.brailnet.cz/sons/docs/zrak/17.htm> >
- [2] KRAUS, Hanuš a kol. *Kompendium očního lékařství*. 1. Vydání. Praha: Grada Publishing, 1997. 341s. ISBN 80-7169-079-1.
- [3] LOBAZ, P. *3D fotografie a alternativní techniky ve fotografii*. Elektronické skriptum. Plzeň: Fakulta aplikovaných věd Západočeské univerzity v Plzni, 2003.
- [4] Effects of age on dynamic accommodation. *Ergonomics* [online]. 2010, **53**(7): 892-903 [cit. 11. listopadu 2015]. DOI: 10.1080/00140139.2010.489968. ISSN 0014-0139. Dostupné z: <http://www.tandfonline.com/doi/abs/10.1080/00140139.2010.489968>
- [5] ATKINSON R. L., ATKINSON R. C., SMITH E. E., BEM D. J., NOLEN-HOEKSEMA S. *Psychologie*. Praha: Portál, 2003. 751 pp. ISBN 80-7178-640-3
- [6] Fyziologie, problémy [online] [cit. 11. listopadu 2015]. Dostupné na www: [http://www.urel.feec.vutbr.cz/web\\_documents/studium/predmety/MVDK/FRVSg/fyziologie.pdf](http://www.urel.feec.vutbr.cz/web_documents/studium/predmety/MVDK/FRVSg/fyziologie.pdf)
- [7] CE ZHU., [ET AL.]. *3D-TV system with depth-image-based rendering architectures, techniques and challenges*. New York, NY: Springer, 2013. ISBN 14-419-9964-7.
- [8] VLACHOVÁ, J. *Stereoskopické promítání*. Praha, 2012. Diplomová práce. Univerzita Karlova v Praze. Vedoucí práce RNDr. Jana Hromadová, Ph.D.
- [9] HOLOTA, R. *Digitalizace a zpracování obrazu*. In FIŘT, Jaroslav; HOLOTA, Radek. *Digitalizace a zpracování obrazu* [online]. Plzeň: Nové technologie - Výzkumné centrum, 2002 [cit. 18. listopadu 2015]. Dostupné z WWW: <http://home.zcu.cz/~holota5/publ/DigZprO.pdf>
- [10] DELIGIANNIDIS, L. *Emerging trends in image processing, computer vision and pattern recognition*. 1st edition. Waltham, MA: Elsevier, 2014, pages cm. ISBN 978-012-8020-456.
- [11] HLAVÁČ V., KYBIC J. *Jasové transformace*. [online]. Praha: Fakulta elektrotechnická ČVUT v Praze, [cit. 20. listopadu 2015]. Dostupné z WWW: <http://cmp.felk.cvut.cz/cmp/courses/33zsl1zima2005/slidy/jasoveTransformace.pdf>
- [12] HORÁK K. *Jasové transformace*. [online]. Brno: Fakulta elektrotechniky a komunikačních technologií VUT, [cit. 20. listopadu 2015]. Dostupné z WWW: [http://midas.uamt.feec.vutbr.cz/ZVS/Lectures/04\\_Jasove\\_transformace.pdf](http://midas.uamt.feec.vutbr.cz/ZVS/Lectures/04_Jasove_transformace.pdf)
- [13] KALOVÁ I. *Předzpracování obrazu*. [online]. Brno: Fakulta elektrotechniky a komunikačních technologií VUT, [cit. 22. listopadu 2015]. Dostupné z WWW: <http://www.uamtold.feec.vutbr.cz/vision/TEACHING/MPOV/02%20-%20Predzpracovani%20obrazu.pdf>

- [14] KLETTE, R. *Concise computer vision: an introduction into theory and algorithms* [online]. 1st edition. Londýn: Springer, 2014, [cit. 29. listopadu 2015]. ISBN 9781447163190.
- [15] STANĚK, Pavel. *Stereoskopický senzor*. Praha, 2007. Diplomová práce. Fakulta elektrotechnická ČVUT v Praze. Vedoucí práce Ing. Miroslav Skrbek, Ph.D.
- [16] PARK J. J. J. H., STOJMENOVIC I., JEONG Y. H., YI G. *Computer science and its applications: ubiquitous information technologies*. 1st edition. New York, NY: Springer Berlin Heidelberg, 2014, pages cm. ISBN 978-366-2454-015.
- [17] VAVROŠ, Ondřej. *Měření vzdálenosti stereoskopickým senzorem*. Brno, 2014. Diplomová práce. Fakulta elektrotechniky a komunikačních technologií VUT. Vedoucí práce Ing. Martin Hasmanda.
- [18] YONG SEOK HEO, KYOUNG MU LEE a SANG UK LEE. Mutual information-based stereo matching combined with SIFT descriptor in log-chromaticity color space. *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2009, , 445-452. DOI: 10.1109/CVPR.2009.5206507. ISBN 978-1-4244-3992-8. Dostupné také z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5206507>
- [19] WANG, Z., A.C. BOVIK, H.R. SHEIKH a E.P. SIMONCELLI. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*. IEEE, 2004, **13**(4), 600-612. DOI: 10.1109/TIP.2003.819861. ISBN 978-1-4244-3992-8. ISSN 1057-7149. Dostupné také z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1284395>
- [20] KLAUS, A., M. SORMANN a K. KARNER. Segment-Based Stereo Matching Using Belief Propagation and a Self-Adapting Dissimilarity Measure. *18th International Conference on Pattern Recognition (ICPR'06)* [online]. IEEE, 2006, , 15-18 [cit. 2. prosince 2015]. DOI: 10.1109/ICPR.2006.1033. ISBN 0-7695-2521-0. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1699458>
- [21] KORDELAS, G.A.; ALEXIADIS, D.S.; DARAS, P.; IZQUIERDO, E. "Revisiting guided image filter based stereo matching and scanline optimization for improved disparity estimation," in *Image Processing (ICIP), 2014 IEEE International Conference on* , vol., no., pp.3803-3807, 27-30 Oct. 2014 doi: 10.1109/ICIP.2014.7025772
- [22] KAMENCAY, P.; BREZNAN, M.; JARINA, R.; LUKAC, P. "Depth map computation using hybrid segmentation algorithm," in *Telecommunications and Signal Processing (TSP), 2011 34th International Conference on* , vol., no., pp.584-588, 18-20 Aug. 2011 doi: 10.1109/TSP.2011.6043662
- [23] FELZENSZWALB, P.F., HUTTENLOCHER, D.P., "Efficient belief propagation for early vision," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on* , vol.1, no., pp.I-261-I-268 Vol.1, 27 June-2 July 2004 doi: 10.1109/CVPR.2004.1315041

- [24] HIRSCHMULLER, H. "Stereo Processing by Semiglobal Matching and Mutual Information," in *Pattern Analysis and Machine Intelligence, IEEE Transactions on* , vol.30, no.2, pp.328-341, Feb. 2008 doi: 10.1109/TPAMI.2007.1166
- [25] BIRCHFIELD, S.; TOMASI, C. "Depth discontinuities by pixel-to-pixel stereo," in *Computer Vision, 1998. Sixth International Conference on* , vol., no., pp.1073-1080, 4-7 Jan 1998 doi: 10.1109/ICCV.1998.710850
- [26] SCHARSTEIN D. *View synthesis using stereo vision*. Berlin: Springer, 1999. ISBN 978-354-0487-258.
- [27] TORR P. H. S., ZISSERMAN A. Feature Based Methods for Structure and Motion Estimation, Proceedings of the International Workshop on *Vision Algorithms: Theory and Practice*, p.278-294, September 21-22, 1999
- [28] JIANG G., Yi J. "Comparison and Study of Classic Feature Point Detection Algorithm," in *Computer Science & Service System (CSSS), 2012 International Conference on* , vol., no., pp.2307-2309, 11-13 Aug. 2012 doi: 10.1109/CSSS.2012.572
- [29] BAY H., ESS A., TUYTELAARS T. aj. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*. 2008, roč. 110, čís. 3, s. 346–359. Dostupné online. ISSN 1077-3142. DOI:10.1016/j.cviu.2007.09.014
- [30] Loopy Belief Propagation, Markov Random Field, Stereo Vision. *Nghiao* [online]. [cit. 2015-12-29]. Dostupné z: [http://nghiaho.com/?page\\_id=1366](http://nghiaho.com/?page_id=1366)
- [31] 3D Vision with Stereo Disparity. *Shawn Lankton Online* [online]. [cit. 2015-12-29]. Dostupné z: <http://www.shawnlankton.com/2007/12/3d-vision-with-stereo-disparity/>
- [32] HIRSCHMÜLLER H., SCHARSTEIN D. Evaluation of cost functions for stereo matching. In IEEE Computer Society Conference on *Computer Vision and Pattern Recognition* (CVPR 2007), Minneapolis, MN, June 2007.
- [33] SEHYUNG L., JIN HAN L., JONGWOO L., IL HONG S. Robust stereo matching using adaptive random walk with restart algorithm, *Image and Vision Computing*, vol. 37, Květen 2015, pp 1-11, ISSN 0262-8856,
- [34] *Mathworks help* [online]. [cit. 2016-05-05]. Dostupné z: <http://www.mathworks.com/help/vision/ref/disparity.html>

## SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

$p_x$	horizontální paralaxa
$f$	ohnisková vzdálenost kamer
$B$	horizontální vzdálenost kamer (stereobase)
$d$	disparita
$Z$	hloubka bodu scény
$x$	horizontální obrazová souřadnice
$y$	vertikální obrazová souřadnice
$I(x, y)$	intenzita obrazového bodu o daných souřadnicích
SSIM	z anglického Structural Similarity Index Measurement
AD	Absolutní difference
$C_{\text{GRAD}}$	podobnostní metrika, rozdíl absolutních hodnot gradientů
MI	vzájemná informace (Mutual information)
3DTV	stereoskopická televize
SIFT	z anglického Scale-Invariant Feature Transform
SURF	z anglického Speeded Up Robust Features